```
A;Status: preliminary, translated from [illegible]/DDBJ
A;Molecule type: mRNA
A;Residues: 1-445 <RES>
A;Cross-references: EMBL:X80198; NID:g931[illegible]; PID:CAA5[illegible]; PID:g95127[illegible]
A;Note: submitted to the EMBL Data Library, July 1994
C;Genetics:
A;Gene: MLN64

  Query Match             25.3%;  Score 37[illegible];  DB 2;  Length 445;
  Best Local Similarity   37.5%;  Pred. No. [illegible]e-21;
  Matches   78;  Conservative   45;  Mismatches   85;  Indels   [illegible];  Gaps   [illegible];

Qy        68 YSDQELAYI[illegible]FAMQFAL[illegible]FVM[illegible]VV[illegible] 12[illegible]
             :| ||  |::||:||         ||: :|  | |     : || | :  |.  ||  : |:   :
Db       231 FSAQEREYIRQGKEATAVVDQILAQEENWKFEKNNEYGDIVYTIEVPFHGKIFILKIILP 2[illegible]

Qy       129 QPMP[illegible]VFRMFAM[illegible]WNPNVK[illegible]AA[illegible] 1[illegible]
              | | :|:|:: : | |  ||  '   ::||::  :|  :::::| | :| |:| |:|
Db       291 [illegible]PAPLVYQPV[illegible]PEEMVLWNKIVTACQ[illegible]YLDVGA[illegible]VV[illegible]RHFVNV [illegible]

Qy       188 RCAKRRGSTCVLACMAIDFCNMPPQKGVIRAH[illegible]CMVLHPLACSP[illegible]KIK[illegible]WL[illegible] 24[illegible]
             |  :||      : :|:||      |       :|  :       ::    | :|         :|: ||
Db       351 RRIERRRDPYLSSCIATSHSAKPPTHKYVR[illegible]CC[illegible]IVLKSACNPPV[illegible]FVW[illegible]NIDL 4[illegible]

Qy       248 KCWLPE[illegible]IN[illegible]VL[illegible]VIFANHLEEEI 27[illegible]
             || ||: :|:| |: |   :|| |||:|:
Db       411 KGRLPRYLIHQSLAAIMFEFAFHLRQRI 43[illegible]


RESULT    6
T22513
hypothetical protein F52F12.7 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T22513
R;Matthews, L.
submitted to the EMBL Data Library, December 1996
A;Reference number: Z19573
A;Accession: T22513
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-241 <WIL>
A;Cross-references: EMBL:Z83228; PIDN:CAB[illegible]5.1; GSPDB:GN[illegible]; CESP:F52F12.7
A;Experimental source: clone F52F12
C;Genetics:
A;Gene: CESP:F52F12.7
A;Map position: 1
A;Introns: 24/3, 125/3, 173/2

  Query Match             12.4%;  Score 1[illegible];  DB 2;  Length 241;
  Best Local Similarity   26.5%;  Pred. No. 4e-07;
  Matches   57;  Conservative   38;  Mismatches   7[illegible];  Indels   17;  Gaps   1[illegible];

Qy        65 EILYSDQELAYI[illegible]FAM[illegible]FAL[illegible]FVM[illegible] 1[illegible]
             | :::|:    ||                     :  |   |:    : || | :'  |    ::  :
Db        40 EAIFNDEN--YL-----------------SHA[illegible]KDESNNEGDVVYAK[illegible]MVI[illegible] 8[illegible]

Qy       125 VVDQPMFPLYFFLVFRMFAMCFWNPNVK[illegible]KTPITH[illegible]AAIN      IV 1[illegible]
              :   |:|  :  :|      |||: |||  |:        ::          :|:             ::
Db        82 ELPMPVEDVMKETWNGMEALPEWNQNINEAA[illegible]AAPTSNEDIVTY        NNEVLVV 1[illegible]

Qy       180 GPRDFVSVRCAKRRGSTCVLACMAIIFCN[illegible]KGVIRAH[illegible]MVL[illegible] 2[illegible]
               :|||  |    ::  |      :||    :         :       :  ||  :||
Db       135 SGREFVSAEIWREV[illegible]FILAS[illegible]SV[illegible]V[illegible] 1[illegible]

Qy       233    SPSKTKLTWLLSIDLKGWLPKSI[illegible]NQV[illegible]TQV 265
                :|   |       ::   |||:|||  |:| |    :   :
Db       187 PENPETTLTDVVMLADLKGYLPKMIVNQV[illegible]M[illegible] 221


RESULT    7
```

```
T16170
hypothetical protein F26F4.4 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 20-Sep-1999 #sequence_revision [illegible]-Sep-[illegible] #text_change 20-Sep-1999
C;Accession: T16170
R;Fulton, L.
submitted to the EMBL Data Library, Mar[illegible]
A;Description: The sequence of C. elegans cosmid F26F4.
A;Reference number: Z18471
A;Accession: T16170
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-478 <FUL>
A;Cross-references: EMBL:[illegible]; NID:[illegible]; PIDN:AA[illegible]; CESP:F26
A;Experimental source: strain Bristol N.
C;Genetics:
A;Gene: CESP:F26F4.4
A;[illegible] 17/[illegible] [illegible]/[illegible] 1[illegible]/[illegible] [illegible] 443/[illegible]

  Query Match              8.9%;  Score [illegible];  DB 2;  Length 478;
  Best Local Similarity   20.2%;  Pred. No. [illegible];
  Matches   62;  Conservative   60;  Mismatches  145;  Indels   40;  Gaps   12;

Qy         1 MLLAT[illegible]KI[illegible]YRHMRNMKGLRQQAV[illegible]IWINQVRRRSS[illegible]IG 6[illegible]
              ::: :| |       |             |      :      :     ' |  '        :          ||:
Db       188 LIIISF[illegible]WSLFY          LMFF[illegible]FERRYAR[illegible]NPEF[illegible]LARSNNRH 2[illegible]9

Qy        61 SRLLE[illegible]YS[illegible]ELAYILQQGLLAMQKA[illegible]NQ[illegible]WKKI[illegible]KVMGK  VVPDVGK 1[illegible]8
               |         |  :               :    :  |           :  :    :      :  :   |:   |  '  :
Db       240 KKGKKQ[illegible]HAPIAVPSKV[illegible]VI[illegible]Y[illegible]AAEF[illegible]PAPC[illegible]VPEDIKP 299

Qy       1[illegible] VPRLEVVV[illegible]LYEELVERMEAMC[illegible]         N[illegible]NVK                 [illegible]KVLQK[illegible]K 1[illegible]2
              :|  |: : : ::::  ||': |    :  |         ' |                         :: :
Db       300 LP  [illegible]                                                        [illegible]

Qy       163 DTPITHELAAEAACNLVGPELPVSV[illegible]VLACMAI[illegible]NMPF[illegible]KG [illegible] VIPA 2[illegible]8
               | : :  :|  |        :   |||: :                     '          |:      '    :|
Db       356 ATDLYYSVSA[illegible]AMRGYIGSRDFLD[illegible]RK[illegible]V[illegible]NNGNEK[illegible]VE[illegible] 4[illegible]5

Qy       21[illegible] [illegible] 278
                :  |:   : :   |          |    ||      |||   |:  : :  :        |:  : :  |: :
Db       416 HNFI[illegible]LK[illegible]EAGVIYF[illegible]FTHEV KKMNEF 469

Qy       279 PASEAKC 285
                :    |
Db       470 AETHYHC 476


RESULT    8
T46357
hypothetical protein DKFZp434G2416.1 - human
C;Species: Homo sapiens (man)
C;Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Feb-2000
C;Accession: T46357
R;Koehrer, K.; Beyer, A.; Mewes, H.W.; [illegible] Wiemann, S.
submitted to the Protein Sequence Database, January 2000
A;Reference number: Z23037
A;Accession: T46357
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-106 <AAA>
A;Cross-references: EMBL:AL137657
A;Experimental source: adult testis; clone DKFZp4[illegible]
C;Genetics:
A;Note: DKFZp434G2416.1

  Query Match              8.1%;  Score 11[illegible];  DB [illegible];  Length 10[illegible];
  Best Local Similarity   28.0%;  Pred. No. [illegible]e-12;
  Matches   28;  Conservative   19;  Mismatches   [illegible];  Indels    1;  Gaps    1;

Qy       178 LVCPRDFVSVRCAKR-RGSTCVLACMA[illegible]FCNMP[illegible]KGVI[illegible]AHH[illegible]CMVLHPLAGSPSK 236
```

```
Db        [illegible] ISERGEV [illegible] VIA [illegible] YFD [illegible] NA [illegible] HVEHE [illegible]

Qy    2[illegible] KLIWL [illegible] DLK [illegible] EK [illegible] INQV [illegible] VDEANH [illegible] PK [illegible]
                 |
Db     6[illegible] NLVI [illegible] NVV [illegible] M [illegible] YAN [illegible] KAV [illegible]


RESULT [illegible]
F6[illegible]25[illegible]
corrinoid methyltransferase protein (mtaC [illegible]) [illegible]
C;Species: Archaeoglobus fulgidus
C;Date: [illegible] Dec [illegible] #sequence_revision [illegible] Dec [illegible] #text_change [illegible]
C;Accession: F6[illegible]
R;Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; [illegible]
 Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; [illegible]
 Glodek, A.; Zh[illegible] [illegible] Weidman [illegible]
Nature [illegible] 364 [illegible]
A;Authors: Utterback, [illegible]
Smith, H.O.; Woese, C.R.; Venter, J.C.
A;Title: The complete genome sequence of the [illegible]
A;Reference number: A6[illegible]; MUID:[illegible]
A;Accession: F6[illegible]
A;Status: preliminary [illegible]
A;Molecule type: DNA
A;Residues: 1-2[illegible]
A;Cross-references: GB:A[illegible]


  Query Match [illegible]
  Best Local Similarity [illegible] Pred. No. [illegible]
  Matches [illegible]56; Conservative [illegible] Mismatches [illegible] Indels [illegible] Gaps [illegible]

Qy      20 MK[illegible]LRQQAVMA[illegible]SQ [illegible] E[illegible] NEEA [illegible] WIN[illegible]EEE [illegible]
           |  :|:  |
Db       1 MADVRKEEV[illegible]ALAD [illegible] EAK[illegible]VE[illegible]KKEVE [illegible] VQKA [illegible]

Qy      72 E[illegible]AYI[illegible]EEAMQKA [illegible] NQE[illegible]WKE [illegible] N[illegible]KVM [illegible]

Db      61 [illegible] MA[illegible]KQV [illegible] EE [illegible] AE [illegible] VVI [illegible]

Qy     123 EVV VI[illegible]MEE[illegible]YE[illegible]VERM[illegible]AM[illegible]EWNENV [illegible]
             |||
Db     116 [illegible]EEVVI[illegible]GVI[illegible]PEA [illegible] VEAANL[illegible]N [illegible] EA [illegible]

Qy     162 KE[illegible]EEI [illegible]PEIA[illegible]AAG [illegible]
             |
Db     169 KGE[illegible]I [illegible] EEA [illegible] 20[illegible]


RESULT 10
G83405
hypothetical protein PA[illegible] [imported] [illegible]
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision [illegible] #text_change [illegible]
C;Accession: G83405
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; M[illegible]
[illegible]dman, S.; Yuan, Y.; Brody, [illegible]
[illegible] Lory, S.; Olson, M.V.
Nature 406, [illegible]-964, 2000
A;Title: Complete genome sequence of [illegible]
A;Reference number: A8[illegible]; MUID:[illegible]
A;Accession: G83405
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-128[illegible]
A;Cross-references: GB:AE0[illegible]; GB:AE[illegible]; NID:[illegible]
A;Experimental source: strain PAO1
C;Genetics:
A;Gene: PA[illegible]


  Query Match [illegible]
```

```
  Best Local Similarity [illegible] Pred. No. [illegible]
  Matches [illegible] Conservative [illegible] Mismatches [illegible] Indels [illegible] Gaps [illegible]

Qy      [illegible] E[illegible]AA[illegible]YE[illegible]MENME[illegible]EQ[illegible]AVMA[illegible]NE [illegible] EAE[illegible]W [illegible]

Db      [illegible] E[illegible]AGN[illegible] [illegible]EKGV[illegible]QAA[illegible]SR[illegible]INGV [illegible] LAVEH [illegible]A[illegible]EAEY [illegible]

Qy      4[illegible] N[illegible]VEEE [illegible] EEEYSE[illegible]EEAY [illegible] AM[illegible]KAE [illegible] N[illegible]E[illegible]WKK [illegible]

Db      [illegible] AAAA[illegible]AG [illegible] W[illegible]AAM [illegible] ANN [illegible]

Qy      [illegible] N[illegible]EKVM[illegible]VV[illegible]V[illegible]KVE [illegible] VVV[illegible]EME [illegible] VEE[illegible]M[illegible]AM [illegible]

Db      [illegible] A[illegible]M [illegible] WQAE [illegible] VMD [illegible] LA [illegible] NE[illegible]AQA[illegible]AEMAEA[illegible]EE [illegible]

Qy      [illegible] KNENVKE[illegible]AV[illegible]K [illegible] AAEAAGNE [illegible] VEE[illegible]ARE [illegible] V[illegible]AAMA [illegible]

Db      [illegible] AAAAGA[illegible]EVAA [illegible] EMA [illegible]

Qy      [illegible] NME[illegible]E [illegible] AEE [illegible] MV[illegible]EE[illegible]AAGE [illegible] EW [illegible] KWEE [illegible]

Db      [illegible] AAA[illegible]A [illegible] EVAEE [illegible] AAAVE[illegible]AEEAA[illegible]AVE [illegible]

Qy      [illegible] N [illegible]

Db      [illegible] VVNE [illegible]


RESULT [illegible]
A[illegible]
probable transposase [illegible]
C;Species: [illegible]
C;Date: 21-[illegible] #sequence_revision [illegible] #text_change 2[illegible] Aug [illegible]
C;Accession: A[illegible]
R;Arthur, M.; M[illegible]; Depardieu, F.; [illegible]
J. Bacteriol. [illegible]
A;Title: Characterization of [illegible] transposon [illegible] glycopeptide
A;Reference number: A[illegible]; MUID:[illegible]
A;Contents: [illegible]
A;Accession: A[illegible]
A;Molecule type: DNA
A;Residues: [illegible] <ART>
A;Cross-references: [illegible] NID:[illegible] PIDN:AAA[illegible]
A;Note: sequence extracted from NCBI backbone [illegible] NCBIP:[illegible]4)
C;Superfamily: [illegible]
C;Keywords: DNA [illegible]


  Query Match [illegible]
  Best Local Similarity [illegible] Pred. No. [illegible]
  Matches [illegible] Conservative [illegible] Mismatches [illegible] Indels [illegible] Gaps [illegible]

Qy      [illegible] W[illegible]N[illegible]VE [illegible]

Db      [illegible] EE[illegible]EEE [illegible] YE[illegible]MEE[illegible]V [illegible] HEN [illegible] YEEYA[illegible]K[illegible]E[illegible]NEE [illegible]

Qy      [illegible] A [illegible] AM[illegible]EA [illegible] N[illegible]E[illegible]WEE [illegible] N[illegible]KVM[illegible]VV [illegible] V [illegible]
                       |
Db      [illegible] KAEEIHE [illegible] K[illegible]EEAL[illegible]E[illegible]A[illegible]N[illegible]KE[illegible]NE[illegible]VIEE[illegible]NE [illegible]

Qy      [illegible] VEE[illegible]EVVV[illegible]EME [illegible] YEE[illegible]VERM[illegible]AM[illegible]EWNENVKEE[illegible]AV [illegible] AAEAA[illegible]NE [illegible]

Db      [illegible] A[illegible]KA[illegible]E [illegible] V[illegible]EVEE[illegible]VEEWN [illegible] V [illegible] VEEA[illegible]E [illegible]

Qy      [illegible] VEE [illegible] V[illegible]V[illegible]AE [illegible] V[illegible]A[illegible]MA [illegible] NMEE[illegible]K[illegible]V [illegible] MV[illegible]EE[illegible]A [illegible] E [illegible]
                 |
Db      [illegible] AE[illegible]A [illegible] Y[illegible]EE[illegible]EEE [illegible] KKY [illegible] V[illegible]EE[illegible]EAN [illegible]

Qy      [illegible] W [illegible] K[illegible]WEE [illegible] INQV [illegible] VDEANH[illegible]PKE [illegible]

Db      [illegible] AV [illegible] MN [illegible] EEEVE [illegible] V [illegible] WE [illegible]
```

```
RESULT  12
C69251
corrinoid methyltransferase protein (mtaC 2 homolog - Archaeoglobus fulgidus
C;Species: Archaeoglobus fulgidus
C;Date: 05 Dec 1997 #sequence_revision [illegible] #text_change [illegible]
C;Accession: C69251
R;Klenk, H.P.; Clayton, R.A.; Tomb, [illegible] Nelson, K.E.; [illegible]
.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.
 Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A;Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.
Smith, H.O.; Woese, C.R.; Venter, J.C.
A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaea
A;Reference number: A69250; MUID:98049343
A;Accession: C69251
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-239 <KLE>
A;Cross-references: GB:AE001106; GB:AE000782; NID:g2689123; PIDN:AAB91208.1; [illegible]

  Query Match             6.9%;  Score 1[illegible];  DB 2;  Length 239;
  Best Local Similarity   28.6%;  Pred. No. [illegible];
  Matches   50;  Conservative   21;  Mismatches   57;  Indels   47;  Gaps   [illegible];

Qy         37 EEALCGPTPSTWINQVREEESLLCGRLEETL[illegible]LQLAYLQQGEEAMQKALCILLNQLLW [illegible]
              :|     |   | |  :    ||: :  ::| |  |       :       :|              | :: :  :   |       |
Db         55 KEVDSCEDPFTILEDVREAIDEIGKRFE[illegible]MACLIIEQVMLF[illegible]

Qy         97 KEESQQQNGEKVMSEV---VPDVCK--VPEL      --EVV     VDQPMHEDLVETLVEPME 142
              ||      :    |    |  :    |        | |:||    |     |           |||       ||||         :|
Db        112 KK--AFSFCFVVIGSVECDVRDICKNIVIALLDDQDTPEVVDICVDQPEEAP          VE 162

Qy        143 AMGEWNPNV----------------KEIKVLQK[illegible]KIEPI -----IHELAAEAA[illegible] 176
              |   :  ||:|                   :  ::  |::  |       |        |  |  |  |    |
Db        163 AANQHNPDVVCLSGLLTEAIESMKKTVEA[illegible] 217

RESULT  13
S54293
regulator protein p122-RhoGAP - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 15-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 01-Sep-199[illegible]
C;Accession: S54293
R;Homma, Y.; Emori, Y.
EMBO J. 14, 286-291, 1995
A;Title: A dual functional signal mediator showing RhoGAP and phospholipase C-delta stim
A;Reference number: S54293; MUID:951[illegible]
A;Accession: S54293
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-1083 <HOM>
A;Cross-references: EMBL:D31962

  Query Match             6.9%;  Score [illegible];  DB 2;  Length 1083;
  Best Local Similarity   20.0%;  Pred. No. [illegible];
  Matches   61;  Conservative   54;  Mismatches   94;  Indels   96;  Gaps   17;

Qy          8 ICAGSSYPHMPNMPGLPQQAVMAISQ[illegible]PSTWINQVREE[illegible] 67
              :|    |   |:   :|   |:  :     :  |      : : :  |                         :
Db        782 VCLAPSLFHLNTLK   PENSSPRVMQ  KA[illegible]KPLQKD---                ENENE [illegible]

Qy         68 YSDQELAYL--------QQGEEAMQ---                ---KALCILLNQ[illegible]       :
               :  |  ||::             |    ||     :                   :|||  |||   :
Db        823 AA[illegible]ARMIADCEE[illegible]

Qy         95 ----------------GWEEPSQQLNE-EVM[illegible]
                                ||          :    :          |       |   :  : : : |       :  |          :   :    :
Db        883 CVDGLEKEVKEEEEGWVSYPDSPQA[illegible]PLPRWPA[illegible]KPLL [illegible]

Qy        139 PEMPAMGEWNPNVKEIKVLQKICEE[illegible]IHELAAEAANLV[illegible]
               :          |:  ::  :  ||::  :         :          :  :|         |            | : |   :    :
```

```
Db        942 KEQHL   WEVDELLDSKVIEELLDSQTEIYQ[illegible]NSMAPHPA        KDYVVLEIWKINL 992

Qy        193 RGSICVLAGMAIDEGNMP   IQKGVIPA[illegible]EVIHPLAG[illegible]KIKLIWEISHCK 248
              ||: [illegible]
Db        993 [illegible]ATVASVEVNVI[illegible] 1044

Qy        249 GWLEK 253
              |  :|:
Db       1045 GHMPE 1049

RESULT  14
E71169
hypothetical protein PH0553 - Pyrococcus horikoshii
C;Species: Pyrococcus horikoshii
C;Date: 14-Aug-1998 #sequence_revision 21-Aug-1998 #text_change 21-[illegible]
C;Accession: E71169
R;Kawarabayasi, Y.; Sawada, M.; Horikawa, H.; [illegible] Yamamoto, S.; Se
 M.; Ohfuku, Y.; Funahashi, T.; Tanaka, [illegible] Kosugi, Y.; Yamazaki, J.; Kushida, N.; Oga
DNA Res. 5, 55-76, 1998
A;Title: Complete sequence and gene organization of the genome of a hyper-thermophili
A;Reference number: A71000; MUID:98344147
A;Accession: E71169
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-[illegible] <KAW>
A;Cross-references: GB:AP000002; NID:g3257129; PIDN:BAA29642.1; PID:g3256959
A;Experimental source: strain OT3
A;Note: this accession replaces an interim accession [illegible] a sequence replaced by GenBa
C;Genetics:
A;Gene: PH0553

  Query Match             6.8%;  Score [illegible];  DB 2;  Length [illegible];
  Best Local Similarity   23.7%;  Pred. No. [illegible];
  Matches   52;  Conservative   36;  Mismatches   77;  Indels   54;  Gaps   11;

Qy         34 ELNKR   ALGGPTPSTWINQVR       EE[illegible]ETL        YSDQELAYLQQGE 80
              |:|  :     |  |    ||   ||::|        :: :|      : :|  :            |    |      |       :
Db         19 EINAKLAE[illegible]KEVIEVVINELREKIMAFE [illegible]AYDSQAAKYEQISK 6[illegible]

Qy         81    EAMQKA[illegible]LLNQ           EGWEEE[illegible]EVVV[illegible] 129
                  :  :   |   ||   ||:|              |     |         :        :           |            |  |:|   : : :
Db         70 RVDELGKK[illegible]SSQLSRLVEAL[illegible]KK[illegible]AVHEKKAES[illegible]AEKAA[illegible]EKVEE[illegible]EEKP 129

Qy        130 ME                  ELYEELVDRM   [illegible]EWN[illegible]VKEIKV [illegible]IHELAAEA [illegible]4
                 |                     :    |||   |::      :  :                 |  :  :  :       |  |    |                |  .  |.
Db        130 [illegible]

Qy        175 AGNLV[illegible]AKRRGQ[illegible]VLA[illegible] 2[illegible]
                  |   |            :    :          ||              |                          :|!::
Db        184 VGELAEKVEEEAEAAKPEEEVVV -          EVEEPE [illegible]

RESULT  15
T05041
hypothetical protein F1[illegible]5.210 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 30-Apr-1999 #sequence_revision [illegible] Apr [illegible] #text_change [illegible] 2000
C;Accession: T05041; T04425
R;Bevan, M.; Pohl, T.; Weitzenegger, [illegible] M.W.[illegible] Mayer, K.F.X.; [illegible]
submitted to the Protein Sequence Database, February 1999
A;Reference number: Z15395
A;Accession: T05041
A;Molecule type: DNA
A;Residues: 1-675 <BEV>
A;Cross-references: EMBL:AL021711
A;Experimental source: [illegible] clone F1[illegible]5
R;Bevan, M.; Bences, V.; Rechmann, S.; [illegible]; Zaccaria, W.; Bancroft, I.; Mewes,
submitted to the Protein Sequence Database, April [illegible]8
A;Reference number: Z15359
A;Accession: T04425
A;Molecule type: DNA
```

```
A;Residues: 1-42[illegible]
A;Cross-references: EMBL [illegible]
A;Experimental [illegible]
C;Genetics:
A;Map position: [illegible]
A;Introns: 16/3 [illegible]
A;Note: [illegible]

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches   50; [illegible]

Qy        30 AISQF NRSA [illegible] IW N VERR [illegible]

Db       166 AFDEEI [illegible] N A  AA  HL [illegible]

Qy        90 L NQF WKE [illegible] KVM FVV [illegible]

Db       207 [illegible] EVL [illegible] BAMKA [illegible] VVE [illegible]

Qy       150 NVKEIEV [illegible] BETAA AA N V [illegible]

Db       249 [illegible] VF [illegible]

Qy       207 NMPE K V  AEH [illegible] MV [illegible]

Db       299 [illegible] N [illegible] K EN [illegible]


Search completed: August [illegible]
Job time: [illegible] secs
```

```
                         GenCore version 4.5
              Copyright (c) 1993 - 200[illegible] Compugen Ltd.

OM protein - protein search, using sw model

Run on:          August 22, 2001, 14:21:[illegible] ; Search time [illegible] Seconds
                                              (without alignments)
                                              164.106 Million cell updates/sec

Title:           US-08-659-254-2
Perfect score:   1429
Sequence:        1 M[illegible]ATFFLCACSSYFHMENM.....   .....VL[illegible] 27[illegible]

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters:      412676

[illegible]mum DB seq length: 0
[illegible]mum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       A_Geneseq_0601:*
                1:  [illegible]AA1990.DAT:*
                2:  [illegible]
                3:  [illegible]
                4:  [illegible]
                5:  [illegible]AA1984.DAT:*
                6:  [illegible]
                7:  [illegible]
                8:  [illegible]
                9:  [illegible]
               10:  [illegible]
               11:  [illegible]
               12:  [illegible]
               13:  [illegible]
               14:  [illegible]
               15:  [illegible]
               16:  [illegible]
               17:  [illegible]
               18:  [illegible]
               19:  [illegible]
               20:  [illegible]
               21:  [illegible]
               22:  [illegible]

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                              SUMMARIES
                    %
Result              Query
  No.     Score     Match Length DB  ID                          Description
----------------------------------------------------------------------------------
    1      1429     100.0    284  20  AAW74153                    Mouse steroidogeni
    2      1429     100.0    284  22  AA[illegible]7.331          Mouse steroidogeni
    3    1258.5      88.1    285  22  AA[illegible]              Human steroidogeni
    4    1252.5      87.6    285  17  AAW[illegible]814           Human steroidogene
    5    1248.5      87.4    285  17  AAW[illegible]              Human steroidogene
    6    1248.5      87.4    285  17  AAW[illegible]              Human steroidogene
    7    1245.5      87.2    285  17  AAW[illegible]237           Human steroidogene
    8    1245.5      87.2    285  17  AAW[illegible]              Human steroidogene
    9    1245.5      87.2    285  17  AAW[illegible]              Human steroidogene
   10      1237      86.6    284  17  AAW[illegible]8             Human steroidogene
   11       374      26.2    445  1[illegible]  AA[illegible]5708  Human MLN [illegible]
   12       312      21.8    180  21  AAB43[illegible]            Human cancer assoc
   13     202.5      14.2     69  21  AAG[illegible]              Human secreted pro
   14       186      13.0    239  21  AAB5[illegible]             Human colon cancer
   15     180.5      12.6    169  21  AAB43[illegible]            Human cancer assoc
   16       177      12.4    221  22  AAB[illegible]              Human steroidogeni
   17       103       7.2     19  20  AAW74[illegible]            Mouse steroid geni
   18       103       7.2     19  22  AAY[illegible]              Mouse tryptic pept
   19       101       7.1    242  21  AAC[illegible]              Arabidopsis thalia
   20       101       7.1    253  21  AAG[illegible]              Arabidopsis thalia
   21      99.5       7.0    291  21  AAB2[illegible]             Protein expressed
   22      99.5       7.0    291  21  AAB2[illegible]             Protein expressed
   23      99.5       7.0    291  22  AAB6[illegible]             Mouse [illegible] protein s
   24      99.5       7.0    291  22  AAB[illegible]              Amino acid sequenc
   25      99.5       7.0    [illegible]  21  AAB[illegible]      Human [illegible] assoc
   26        99       6.9    242  21  AAG[illegible]              Arabidopsis thalia
   27        99       6.9    242  21  AAG[illegible]              Arabidopsis thalia
   28      96.5       6.8   2763  18  AAW[illegible]44            Maize dwarf mosaic
   29      93.5       6.5   1091  20  AAY[illegible]              Protein encoded by
   30      92.5       6.5   1605  21  AAY[illegible]              Amino acid sequenc
   31        92       6.4    257  21  AAY[illegible]              F-box protein FBP
   32        92       6.4   1509  21  AAY41[illegible]8           Amino acid sequenc
   33        89       6.2    389  21  AAB[illegible]              Human [illegible] ORF1324
   34        89       6.2    598  21  AAY[illegible]              Human Goodpasture
   35        87       6.1    598  21  AAY[illegible]              Cattle Goodpasture
   36      86.5       6.1   1517  21  AAY[illegible]26            Amino acid sequenc
   37      86.5       6.1   1588  21  AAY[illegible]              Amino acid sequenc
   38      85.5       6.0    456  21  AAG[illegible]              Arabidopsis thalia
   39      85.5       6.0    484  21  AAG[illegible]              Arabidopsis thalia
   40      85.5       6.0    518  21  AAG[illegible]              Arabidopsis thalia
   41      85.5       6.0    596  21  AA[illegible]              Goodpasture antige
   42      85.5       6.0    616  21  AA[illegible]              Goodpasture antige
   4[illegible] [illegible]                                      Goodpasture antige
   44      [illegible]                                          [illegible]
   45      [illegible]                     21  AAY[illegible]    Human Goodpasture

                              ALIGNMENTS

RESULT 1
AAW7415[illegible]
ID   AAW7415[illegible] standard; Protein; 284 AA.
XX
AC   AAW7415[illegible];
XX
DT   [illegible]-MAY-199[illegible]  (first entry)
XX
DE   Mouse steroidogenic acute regulatory protein
XX
KW   [illegible] McCune-Albright syndrome;
KW   steroid hormone dependent disorder; adrenal hypoplasia congenita; LCAH;
KW   lipoid [illegible] adrenal [illegible] maturation;
KW   androgen responsive tumours; hypogonadism; therapy; progesterone level.
XX
OS   Mus sp.
XX
PN   US587224[illegible]-A.
XX
PD   16-FEB-199[illegible].
XX
PF   04-NOV-1995;   95US-05389[illegible].
XX
PR   04-NOV-1995;   95US-05389[illegible].
XX
PA   (UYTE-) UNIV TEXAS TECH HEALTH SCI CENT.
XX
PI   Clark BJ,  Stocco DM;
XX
DR   WPI; 199[illegible]-166726/14.
DR   N-PSDB; AAX18169, AAX1817[illegible].
XX
PT   New nucleic acid encoding steroidogenic acute regulatory protein
PT   useful for treatment of steroid hormone dependent disorders such as
```

```
FT   lipoid congenital adrenal hyperplasia [illegible]
FT   androgen response [illegible]
XX
PS   Claim 1; [illegible]
XX
CC   This sequence [illegible] the steroidogenic [illegible]
CC   of the invention. [illegible] StAR [illegible]
CC   derivatives may be used as probes and primers [illegible]
CC   studies and in the treatment [illegible]
CC   as lipoid congenital adrenal hyperplasia [illegible]
CC   maturation and androgen response [illegible]
CC   McCune-Albright syndrome [illegible]
CC   Monitoring levels of StAR [illegible]
CC   and so [illegible]
XX
SQ   Sequence [illegible]

  [illegible]ery Match [illegible]
  [illegible]est Local Similarity [illegible]
  Matches [illegible]; Conservative [illegible] Mismatches [illegible] Gaps [illegible]

Qy          1 MFLA[illegible]
Db          1 mfla[illegible]

Qy         61 QLEA[illegible]EVAM[illegible]NN[illegible]WKK[illegible]
Db         61 qlea[illegible]

Qy        121 [illegible]EVVV[illegible]WNENVK[illegible]
Db        121 [illegible]

Qy        181 [illegible]
Db        181 [illegible]

Qy        241 [illegible]
Db        241 [illegible]


RESULT [illegible]
AAY72831
[illegible] AAY72831 standard; [illegible] AA.
[illegible]
[illegible] AAY72831;
XX
DT   31-MAY-2001  (first entry)
XX
DE   Mouse steroidogenic acute regulatory [illegible]
XX
KW   Mouse; steroidogenic acute regulatory [illegible]
KW   gene therapy; steroid hormone dependent disorder; [illegible]
KW   lipoid congenital adrenal hyperplasia; [illegible]
KW   infertility; androgen response [illegible]
KW   hypogonadotropic hypogonadism; McCune-Albright [illegible]
KW   adrenal hyperplasia [illegible]
XX
OS   Mus sp.
XX
FH   Key                 Location/Qualifiers
FT   Peptide             [illegible]
FT                       /label= [illegible]
FT   Cleavage-site       [illegible]
FT                       /note= [illegible]
FT                       [illegible]
FT                       [illegible]
FT   Protein             [illegible]
FT                       /label= Mature [illegible]
FT   Cleavage-site       [illegible]
FT                       /note= [illegible]
```

```
[illegible]
XX
[illegible]
XX
[illegible]
XX
CC   [illegible] StAR [illegible]
CC   [illegible] therapy for the treatment of steroid
CC   hormone dependent disorders [illegible]
CC   adrenal hyperplasia [illegible] infertility [illegible]
CC   [illegible] McCune-Albright
CC   [illegible]
CC   [illegible] StAR
XX
SQ   Sequence [illegible] AA;

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches [illegible] Mismatches [illegible] Gaps [illegible]

Qy          1 [illegible]
Db          1 [illegible]

Qy         61 [illegible]
Db         61 [illegible]

Qy        121 [illegible]
```

```
             ||||||||||||:||||||||||||||||||| | [illegible]
Db     121 [illegible]

Qy     181 RDFVSVRCTKRRGSTCVLAGMATHFGEMP[illegible] 24
             :||||||||||||||||||||||||||||| [illegible]
Db     181 rdfvsvrctkrrgstcvlagmat[illegible] 24

Qy     241 LLSIDLKGWLPKFIINQVLSQTQIEFANHLRK [illegible] 27
             |||||||||||||||||||||||||||||| [illegible]
Db     241 llsidlkgwlpk[illegible] 27


RESULT    3
AAY72838
ID   AAY72838 standard; Protein; 285 AA.
XX
AC   AAY72838;
XX
DT   31-MAY-2001  (first entry)
XX
DE   Human steroidogenic acute regulatory (StAR) protein.
XX
KW   Human; steroidogenic acute regulatory [illegible] StAR;
KW   gene therapy; steroid hormone-dependent disorder; LCAH;
KW   lipoid congenital adrenal hyperplasia; sexual maturation;
KW   infertility, androgen responsive tumour; precocious puberty;
KW   hypogonadotropic hypogonadism; McCune-Albright syndrome;
KW   adrenal-hypoplasia congenita; cytostatic.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 156
FT                   /note= "Encoded by CAC, this feature corresponds
FT                   to the translation exception for the sequence shown
FT                   in (AAD02909)"
FT   Modified-site   59..60
FT                   /note= "PKA phosphorylation site"
FT   Modified-site   68..69
FT                   /note= "CK phosphorylation site"
FT   Modified-site   195
FT                   /note= "PKA phosphorylation site"
FT   Modified-site   233..234
FT                   /note= "P34CDC2 kinase (Cb) phosphorylation site"
FT   Modified site   238
FT                   /note= "PKC phosphorylation site"
FT   Modified-site   277
FT                   /note= "PKA phosphorylation site"
XX
PN   US6194555-B1.
XX
     27-FEB-2001.
XX
PF   07-JUN-1996;   96US-0659254.
XX
PR   04-NOV-1995;   95US-0538960.
XX
PA   (UYTE-) UNIV TEXAS TECH HEALTH S[illegible]N.
XX
PI   Stocco DM,  Clark BJ;
XX
DR   WPI; 2001-217936/22.
DR   N-PSDB; AAD02909, AAD02912.
XX
PT   Novel nucleic acid encoding steroidogenic acute regulatory protein,
PT   useful for treating steroid hormone dependent disorders -
XX
PS   Example 14; Fig 5; 41pp; English.
XX
CC   The invention relates to steroidogenic acute regulatory protein
CC   (StAR) and nucleic acid molecules encoding them. StAR proteins
CC   are used for regulating steroidogenesis. More particularly, these
CC   proteins are involved in regulating cholesterol transport into
```

```
CC   the mitochondria of a cell. The nucleic acid molecules of the
CC   invention are useful in gene therapy for the treatment of steroid
CC   hormone-dependent disorders, particularly, lipoid congenital
CC   adrenal hyperplasia (LCAH), infertility, sexual maturation,
CC   androgen-responsive tumours, precocious puberty, McCune Albright
CC   syndrome, adrenal hypoplasia congenita or hypogonadotropic
CC   hypogonadism. They are also useful for recombinant preparation
CC   of the steroidogenic acute regulatory protein.
CC   The present sequence is human steroidogenic acute regulatory
CC   protein (StAR).
XX
SQ   Sequence   285 AA;

  Query Match             88.1%;  Score [illegible];  DB 22;  Length 285;
  Best Local Similarity   87.4%;  Pred. No. [illegible] 121;
  Matches  242;  Conservative   18;  Mismatches  [illegible];  Indels  [illegible];  Gaps  [illegible];

Qy       1 MFLATFKL[illegible]AG[illegible]YRHMRNMKGLRHQAVLA[illegible]WMGQVRRRSSL[illegible] 59
             | ||||:|||||||||||||||||| ||||:| [illegible]
Db       1 mllattk[illegible]syrhmrnmkglrqq[illegible]rss [illegible] 6

Qy      60 [illegible]LLA[illegible]VAMQKAL[illegible]KVL 11
             |:|| | [illegible]
Db      61 s[illegible] 12

Qy     120 RLEVVVDQPMERLYEELVERMEAMG[illegible]WNPNVKE[illegible]KV[illegible] 17
             |||||| [illegible]
Db     121 rlevvvdqpm[illegible] 180

Qy     180 PRDFVSVR[illegible]VLAGMA[illegible]KLL 23
             |||||||| ||||  ||||||||| | [illegible]
Db     181 prdfvsvr[illegible]sktkll 240

Qy     240 WLLSIDLKGWLPKTIINQVLSQTQIEFAN[illegible]A 27
             ||||||||:|||||:|||||||||||: || | [illegible]
Db     241 wllsidlkqwipksiinqvlsqt[illegible] 27


RESULT    4
AAW00814
ID   AAW00814 standard; Protein; 285 AA
XX
AC   AAW00814;
XX
DT   21-MAY-1997  (first entry)
XX
DE   Human steroidogenesis acute regulatory protein.
XX
KW   Human; steroidogenesis; acute regulatory protein; hStAR; analysis;
KW   mutation; detection; prenatal genetic defect; [illegible] protein;
KW   lipoid adrenal hyperplasia; treatment; prevention; gene;
KW   replacement therapy; hypercholesterolaemia
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Region          5
FT                   /note= "potential site for protein kinase C
FT                          mediated phosphorylation"
FT   Region          13
FT                   /note= "potential site for protein kinase C
FT                          mediated phosphorylation"
FT   Region          56..57
FT                   /note= "potential site for protein kinase A
FT                          mediated phosphorylation"
FT   Region          100
FT                   /note= "potential site for protein kinase A
FT                          mediated phosphorylation"
FT   Region          1[illegible]
FT                   /note= "potential site for protein kinase A
FT                          mediated phosphorylation"
FT   Region          186
FT                   /note= "potential [illegible]
FT                          mediated phosphorylation [illegible]"
FT   Region          [illegible]
FT                   /note= "potential [illegible]
FT                          mediated phosphorylation [illegible]"
FT   Region          [illegible]
FT                   /note= "potential [illegible]
FT                          mediated phosphorylation [illegible]"
XX
PN   WO9629338-A1.
XX
PD   26-SEP-1996.
XX
PF   22-MAR-1996;   96WO-US03906.
XX
PR   23-MAR-1995;   95US-[illegible]
XX
     (REGC ) UNIV CALIFORNIA.
     (UYPE ) UNIV PENNSYLVANIA.
XX
PI   Lin D,  Miller WL,  Strauss JF;
XX
DR   WPI; 1996-443[illegible]/44.
DR   N-PSDB; AA[illegible]
XX
PT   Isolated human steroidogenic acute regulatory [illegible]
PT   for detection of mutation(s) [illegible]
PT   lipoid adrenal hyperplasia
XX
PS   Claim 1; Page [illegible]
XX
CC   The present sequence is the human steroidogenic acute
CC   regulatory protein (StAR). The StAR [illegible]
CC   mutations to detect [illegible]
CC   with congenital lipoid adrenal hyperplasia (CAH [illegible]
CC   transmission [illegible] CAH [illegible]
CC   replacement therapy, which can also be used [illegible]
CC   hypercholesterolemia
CC   A human adrenal cortex cDNA library was [illegible]
CC   probe to isolate [illegible]
CC   residue protein. When it was [illegible]
CC   COS-1 cells [illegible]
CC   level of pregnenolone [illegible]
CC   20 alpha hydroxycholesterol.
XX
SQ   Sequence   285 AA;

  Query Match             [illegible]
  Best Local Similarity   [illegible]
  Matches  [illegible]  Mismatches [illegible]

Qy          1 MLLATFKLCAG[illegible]NMK[illegible]
              [illegible]
Db          1 mllatfkl[illegible]

Qy         60 SQLEATLY[illegible]VAMQKA[illegible]NN[illegible]
              [illegible]
Db         61 sqleetly[illegible]

Qy        120 RLFVVVVD[illegible]
              [illegible]
Db        121 r[illegible]

Qy        180 PRGFVSVR[illegible]
              [illegible]
Db        181 prg[illegible]

Qy        240 WL[illegible]DLK[illegible]
              [illegible]
Db        241 wl[illegible]dlk[illegible]
```

```
RESULT [illegible]
AAW[illegible]
ID   AAW[illegible] standard; Protein; 285 AA.
XX
AC   AAW[illegible]
XX
DT   [illegible] MAY [illegible] (first entry)
XX
DE   Human steroidogenic acute regulatory [illegible] mutant A[illegible]V[illegible]
XX
KW   [illegible] StAR [illegible]
KW   mutations [illegible] detect [illegible] genetic [illegible]
KW   lipoid adrenal hyperplasia; treatment; [illegible] gene
KW   replacement therapy; [illegible]
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference [illegible]
FT                   /note= "[illegible] A[illegible] with Val"
XX
PN   W[illegible]
XX
PD   [illegible]
XX
PF   [illegible] MAR [illegible]
XX
PR   [illegible] MAR [illegible]
XX
PA   (REGC ) UNIV CALIFORNIA.
PA   (UYPE ) UNIV PENNSYLVANIA.
XX
PI   Lin D,  Miller WL,  Strauss JF;
XX
DR   WPI; [illegible]
DR   N-PSDB; AA[illegible]
XX
PT   [illegible]
PT   for detection of mutation(s) [illegible] this gene that cause congenital
PT   lipoid adrenal hyperplasia
XX
PS   Claim [illegible]
XX
CC   The present sequence is the human steroidogenic acute regulatory
CC   protein (StAR) [illegible] lipoid adrenal hyperplasia (CAH)
CC   [illegible] A[illegible]V[illegible] Analysing the StAR gene for [illegible]
CC   mutation [illegible] CAH
CC   associated genetic defects, or its transmission to children. CAH
CC   [illegible] gene replacement therapy, which can
CC   also be used [illegible]
XX
SQ   Sequence   285 AA;

  Query Match             [illegible]
  Best Local Similarity   [illegible]
  Matches [illegible]  Conservative [illegible]  Mismatches [illegible]  Indels [illegible]  Gaps [illegible]

Qy          1 MLL[illegible]MENMK[illegible]WM[illegible]VRRR[illegible]
              [illegible]
Db          1 [illegible]

Qy         6[illegible] SQLE[illegible]VAMQKA[illegible]NN[illegible]PVW[illegible]KMVPKVGKV[illegible]
              [illegible]
Db         6[illegible] [illegible]

Qy        1[illegible] [illegible]
              [illegible]
Db        1[illegible] [illegible]

Qy        1[illegible] [illegible]
              [illegible]
Db        1[illegible] [illegible]
```

```
Qy     240 WLLSIDLKGWLPKTIINQVLSQTQIFFANHLRKRLEA 276
           ||||||||||||:|||||||||||::||| ||  |:
Db     241 wllsidlkgwlpksii[illegible] 277


RESULT    6
AAW13235
ID    AAW13235 standard; Protein; 285 AA.
XX
AC    AAW13235;
XX
DT    21-MAY-1997  (first entry)
XX
DE    Human steroidogenesis acute regulatory protein mutant Glu169Lys.
XX
KW    Human; steroidogenesis; acute regulatory protein; hStAR; analysis;
KW    mutation; detection; prenatal; genetic defect; congenital; protein;
KW    lipoid adrenal hyperplasia; treatment; prevention; gene;
KW    replacement therapy; hypercholesterolaemia; mutant.
XX
      Homo sapiens.

      Key               Location/Qualifiers
FT    Misc-difference   169
FT                      /note= "wild type Glu substituted with Lys"
XX
PN    WO9629338-A1.
XX
PD    26-SEP-1996.
XX
PF    22-MAR-1996;   96WO-US03896.
XX
PR    23-MAR-1995;   95US-0410540.
XX
PA    (REGC ) UNIV CALIFORNIA.
PA    (UYPE-) UNIV PENNSYLVANIA.
XX
PI    Lin D,  Miller WL,  Strauss JF;
XX
DR    WPI; 1996-443130/44.
DR    N-PSDB; AAT62020.
XX
PT    Isolated human steroidogenesis acute regulatory protein gene - used
PT    for detection of mutation(s) of this gene that cause congenital
PT    lipoid adrenal hyperplasia
XX
PS    Claim 13; Page -; 89pp; English.
XX
      The present sequence is the human steroidogenesis acute regulatory
      protein (hStAR), congenital lipoid adrenal hyperplasia (CAH)
      associated mutant, Glu169Lys. Analysis of the hStAR gene for such a
      mutation is useful in the detection (e.g. prenatally) of CAH
CC    associated genetic defects, or its transmission to children. CAH
CC    can be treated by protein or gene replacement therapy, which can
CC    also be used to prevent or treat hypercholesterolaemia.
XX
SQ    Sequence   285 AA;


  Query Match              87.4%;  Score [illegible];  DB 17;  Length 285;
  Best Local Similarity    86.6%;  Pred. No. [illegible];
  Matches  240;  Conservative   19;  Mismatches   17;  Indels   [illegible];  Gaps   1;

Qy       1 MFLAIFKL[illegible]
           | ||||||||||||||||||||||||| | :::    : | ||||  :|   |: ||    : :
Db       1 mllatfkl[illegible]

Qy      60 SQLFATLYSDQELSY[illegible]
           |:|| ||||:|||:|:||||  :||||  |:     |  |:  ||||::||:|:|  : :   |:|
Db      61 srleetlysdqelayl[illegible]

Qy     120 R[illegible]FVVVPQPMERLYEELVERMEAMGEWNP[illegible]
```

```
           ||||||||||:|:|||||:||||||||   |  || : :||||      : : | | : : |
Db     121 [illegible]

Qy     180 [illegible] 239
           |||||||||  :||||||||||||  |  ||  |     :|  : : :  |  |||  ||||||||:||  |
Db     181 prdfvsvrcakrrgstcvlagmatdf[illegible] 240

Qy     240 WLLSIDLKGWLPKTIINQVLSQTQIEFANH[illegible] 276
           ||||||||||:|:|||||||||||::||| | ||  |:
Db     241 wllsidlkgwlpksiinqvlsqtqvdfan[illegible] 277


RESULT    7
AAW13237
ID    AAW13237 standard; Protein; 285 AA.
XX
AC    AAW13237;
XX
DT    21-MAY-1997  (first entry)
XX
DE    Human steroidogenesis acute regulatory protein mutant Leu275Pro.
XX
KW    Human; steroidogenesis; acute regulatory protein; hStAR; analysis;
KW    mutation; detection; prenatal; genetic defect; congenital; protein;
KW    lipoid adrenal hyperplasia; treatment; prevention; gene;
KW    replacement therapy; hypercholesterolaemia; mutant.
XX
OS    Homo sapiens.
XX
FH    Key               Location/Qualifiers
FT    Misc-difference   275
FT                      /note= "wild type Leu substituted with Pro"
XX
PN    WO9629338-A1.
XX
PD    26-SEP-1996.
XX
PF    22-MAR-1996;   96WO-US03896.
XX
PR    23-MAR-1995;   95US-0410540.
XX
PA    (REGC ) UNIV CALIFORNIA.
PA    (UYPE-) UNIV PENNSYLVANIA.
XX
PI    Lin D,  Miller WL,  Strauss JF;
XX
DR    WPI; 1996-443130/44.
DR    N-PSDB; AAT62022.
XX
PT    Isolated human steroidogenesis acute regulatory protein gene - used
PT    for detection of mutation(s) of this gene that cause congenital
PT    lipoid adrenal hyperplasia
XX
PS    Claim 13; Page -; 89pp; English.
XX
CC    The present sequence is the human steroidogenesis acute regulatory
CC    protein (hStAR), congenital lipoid adrenal hyperplasia (CAH)
CC    associated mutant, Leu275Pro. Analysis of the hStAR gene for such a
CC    mutation is useful in the detection (e.g. prenatally) of CAH
CC    associated genetic defects, or its transmission to children. CAH
CC    can be treated by protein or gene replacement therapy, which can
CC    also be used to prevent or treat hypercholesterolaemia.
XX
SQ    Sequence   285 AA;


  Query Match              87.2%;  Score [illegible];  DB [illegible];  Length 285;
  Best Local Similarity    86.6%;  Pred. No. [illegible];
  Matches  24[illegible];  Conservative   18;  Mismatches   [illegible];  Indels   1;  Gaps   1;

Qy       1 MFLAIFKL[illegible]
           | ||||||  ||||||||||||||     : |  :   |:    |   ::  : ||||||
Db       1 mllatfkl[illegible]patwinqvrrrsslg 60
```

```
Qy       60 [illegible]VAM[illegible]KA[illegible]NN[illegible]WKK[illegible]
              [illegible]
Db       61 [illegible]
 
Qy      120 [illegible]EVVV[illegible]EMFRIY[illegible]ELVDRMEAM[illegible]WNPNVKE[illegible]
              [illegible]
Db      121 [illegible]
 
Qy      180 [illegible]DEV[illegible]V[illegible]AGMA[illegible]EMP[illegible]
              [illegible]
Db      181 [illegible]
 
Qy      240 W[illegible]LSI[illegible]W[illegible]PKI[illegible]NQVI[illegible]ANHKKR[illegible]A
              [illegible]
Db      241 [illegible]
 
 
[illegible]T   [illegible]
[illegible]3233
ID   AAW[illegible] standard; [illegible]; AA.
XX
AC   AAW[illegible];
XX
DT   21 MAY 1997  (first entry)
XX
DE   Human [illegible]
XX
KW   Human; steroidogenesis; acute regulatory; [illegible]
KW   mutation; detection; prenatal; [illegible]
KW   lipoid adrenal hyperplasia; [illegible]
KW   replacement therapy; [illegible]
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference [illegible]
FT                   /note= "wild type [illegible]"
XX
PN   WO9629[illegible]-A1.
XX
PD   26-SEP-1996.
XX
PF   22-MAR-1996;   96WO-[illegible]
XX
PR   23-MAR-1995;   [illegible]
XX
PA   (REGC ) UNIV CALIFORNIA.
PA   (UYPE ) UNIV PENNSYLVANIA.
XX
PI   Lin D,  Miller W, [illegible]
XX
DR   WPI; 1996-44[illegible]
DR   N-PSDB; AAT[illegible]
XX
PT   Isolated human [illegible]
PT   for detection [illegible]
PT   lipoid adrenal hyperplasia
XX
PS   Claim [illegible]; Page [illegible]
XX
CC   The present sequence [illegible] human [illegible]
CC   protein (hStAR) [illegible] lipoid adrenal hyperplasia [illegible]
CC   associated mutant [illegible]
CC   mutation is useful in the detection [illegible]
CC   associated genetic defects [illegible]
CC   can be treated by [illegible]
CC   is be used [illegible]
XX
SQ   Sequence   285 AA;
 
  Query Match [illegible]
```

```
[illegible]
[illegible] Matches [illegible]
 
Qy   [illegible]
Db   [illegible]
 
Qy   [illegible]VAM[illegible]KA[illegible]NN[illegible]WKK[illegible]
Db   [illegible]
 
Qy   [illegible]VVV[illegible]EMFRIY[illegible]ELVDRMEAM[illegible]WNPNVKE[illegible]
Db   [illegible]
 
Qy   [illegible]
Db   [illegible]
 
Qy   [illegible]NQVI[illegible]ANHKKR[illegible]
Db   [illegible]
 
 
[illegible]
AAW[illegible]
ID   AAW[illegible] standard; [illegible]; AA.
XX
AC   AAW[illegible];
XX
DT   [illegible] MAY [illegible] (first entry)
XX
DE   Human [illegible] mutant A[illegible]
XX
KW   [illegible] hStAR [illegible]
KW   [illegible]
KW   [illegible]
KW   [illegible]
XX
OS   [illegible]
XX
FH   Key             Location/Qualifiers
FT   Misc-difference [illegible]
FT                   /note= "wild type A[illegible] with [illegible]"
XX
PN   W[illegible]-A1.
XX
PD   [illegible]
XX
PF   [illegible] MAR [illegible]  [illegible]WO-[illegible]
XX
PR   [illegible] MAR [illegible]
XX
PA   [illegible] UNIV CALIFORNIA.
PA   [illegible] UNIV PENNSYLVANIA.
XX
PI   [illegible] Miller W[illegible]
XX
DR   WPI; [illegible]
DR   N-PSDB; AAT[illegible]
XX
PT   [illegible]
PT   [illegible]
PT   [illegible]
XX
PS   [illegible]
XX
CC   The present sequence [illegible]
CC   [illegible] (hStAR) [illegible] lipoid adrenal hyperplasia (LAH)
CC   [illegible] hStAR [illegible]
CC   [illegible] LAH
CC   associated genetic defects [illegible]
CC   can be treated by [illegible] replacement therapy [illegible]
CC   also be used to prevent or treat hyper[illegible]lesterolaemia.
XX
SQ   Sequence   285 AA;

  Query Match             87.2%;  Score 124[illegible];  DB 17;  Length 285;
  Best Local Similarity   86.6%;  Pred. No. [illegible]e-120;
  Matches  240;  Conservative   18;  Mismatches   18;  Indels   [illegible];  Gaps   [illegible];

Qy        1 MFLATFKLCAGSSYRHMRNMK[illegible]LRHQAVLA[illegible] [illegible]
Db        1 mllattk[illegible]rqgsyrhmrnmkglrqqavmal[illegible] [illegible]

Qy       60 SQLFATLYSDQ[illegible]SYL[illegible]EVAMQKA[illegible]NN[illegible] [illegible]
Db       61 srleetlysdqe[illegible] [illegible]

Qy      120 RLFVVVDQPMDRL[illegible]FFLVDRMEAMGEWNLNVK[illegible] [illegible]
Db      121 rlevvvdqp[illegible] [illegible]

Qy      180 PRDFVSVRCTKRRGSTCVLAGMATHFGRMPP[illegible] [illegible]
Db      181 pldfvsvrcak[illegible]gst[illegible]vlagm[illegible]dfg[illegible] [illegible]

Qy      240 WLLSIDLKGWLPKTIINQVLSQTQIEFANHLRK[illegible]PA 276
            ||||||||||| :||||||||||||::|| |||  |||:
Db      241 wllsidlkgwlpksiinqvlsqtqvdfa[illegible] [illegible]


RESULT  10
AAW13238
ID   AAW13238 standard; Protein; 284 AA.
XX
AC   AAW13238;
XX
DT   21-MAY-1997  (first entry)
XX
DE   Human steroidogenesis acute regulatory protein mutant delta Arg272.
XX
KW   Human; steroidogenesis; acute regulatory protein; hStAR; analysis;
KW   mutation; detection; prenatal; genetic defect; congenital protein;
KW   lipoid adrenal hyperplasia; treatment; prevention; gene;
KW   replacement therapy; hypercholesterolaemia; mutant.
XX
OS   Homo sapiens.
XX
PN   WO9629338-A1.
XX
PD   26-SEP-1996.
XX
PF   22-MAR-1996;   96WO-US03896.
XX
PR   23-MAR-1995;   95US-0410540.
XX
PA   (REGC ) UNIV CALIFORNIA.
PA   (UYPE-) UNIV PENNSYLVANIA.
XX
PI   Lin D,  Miller WL,  Strauss JF;
XX
DR   WPI; 1996-443130/44.
XX
PT   Isolated human steroidogenesis acute regulatory protein gene - used
PT   for detection of mutation(s) of this gene that cause congenital
PT   lipoid adrenal hyperplasia
XX
PS   Claim 16; Page -; 89pp; English.
XX
CC   The present sequence is the human steroidogenesis acute regulatory
CC   protein (hStAR), congenital lipoid adrenal hyperplasia (CAH)
CC   associated mutant, delta Arg272 (i.e. wild type residue Arg272 has
CC   been deleted). Analysing the hStAR gene for such a mutation is
CC   useful in the detection (e.g. prenatally) of CAH associated
```

```
CC   genetic defects, or its transmission to children. CAH can be
CC   treated by protein or gene replacement therapy, which can also be
CC   used to prevent or treat hypercholesterolaemia
XX
SQ   Sequence   284 AA;

  Query Match             86.6%;  Score [illegible];  DB 17;  Length 284;
  Best Local Similarity   86.6%;  Pred. No. [illegible]e-119;
  Matches  240;  Conservative   18;  Mismatches   17;  Indels    2;  Gaps    2;

Qy        1 MFLATFKLCAGSSYRHMRNMKGLRHQAVLA[illegible]WNQVRRRSSL[illegible] 59
Db        1 mllattkl[illegible]rqgsyrhmrnmkglrqqavmal[illegible]rrrss[illegible] 60

Qy       60 SQLFAT[illegible]EVAMQKA[illegible]DVGKVF [illegible]
Db       61 srleetlysd[illegible]aylqqaeeamqka[illegible]kvmakvqp[illegible] [illegible]

Qy      120 RLFVVV[illegible]MEAMGEWN[illegible]NLVG [illegible]
Db      121 rlevvvdqp[illegible]lyse[illegible]rmeamqewnp[illegible]qkiqkdt[illegible] 180

Qy      180 PRDFVSVR[illegible]VLAGMATHFG[illegible]KIRLL 239
Db      181 prdfvsv[illegible]qstcvlagmatdfgnm[illegible]qkq[illegible]plaaspskt[illegible] 240

Qy      240 WLLSIDLKGWLPKTIINQVLSQTQIEFANHLRKRLE[illegible]A 276
Db      241 wllsidlkgwlpksiinqvlsqtqvdfa[illegible]kr[illegible]s 276


RESULT  11
AAW25768
ID   AAW25768 standard; Protein; 445 AA.
XX
AC   AAW25768;
XX
DT   05-DEC-1997  (first entry)
XX
DE   Human MLN 64.
XX
KW   Human; CAB1; Casp 1; MLN 64; MLN [illegible]; [illegible]; breast carcinoma cell;
KW   Cys rich region, RING finger domain; CAB1 motif; [illegible] domain; LIM domain;
KW   TNF-related cytokine signal transduction; breast cancer; SH3;
KW   Src homology region 3.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Domain          1..72
FT                   /label= Potential transmembrane domain
FT   Domain          94..168
FT                   /label= Potential transmembrane domain
XX
PN   WO9706256-A2.
XX
PD   20-FEB-1997.
XX
PF   31-JUL-1996;   96WO-US12500.
XX
PR   09-AUG-1995;   95US-0002183.
XX
PA   (BRIM ) BRISTOL-MYERS SQUIBB CO.
PA   (CNRS ) CENT NAT RECH SCI.
PA   (INRM ) INST NAT SANTE & RECH MEDICALE.
PA   (UYPA-) UNIV PASTEUR LOUIS.
XX
PI   Basset P,  Byrne J,  Rio M,  Tomasetto C;
XX
DR   WPI; 1997-154263/14.
DR   N-PSDB; AAT86171.
XX
```

```
PT   Gene amplif[illegible] over expre[illegible]
PT   chromosome [illegible]
PT   prognosis and [illegible] markers
XX
PS   Claim 11; [illegible]
XX
CC   This sequence [illegible] MLN 64 [illegible]
CC   the MLN 64 gene has been mapped [illegible]
CC   of chromosome [illegible] with the max[illegible]
CC   contains two [illegible]
CC   gene contains [illegible] There are [illegible] MLN [illegible]
CC   from [illegible]
CC   modification [illegible]
CC   the MLN 64 variants [illegible]
CC   variants lead [illegible]
CC   C truncated MLN 64 [illegible]
CC   containing [illegible]
CC   residues. The full length MLN [illegible]
CC   transmembrane [illegible] and [illegible]
CC   leucine rich repeat [illegible] MLN [illegible]
CC   useful as [illegible] markers [illegible]
XX
SQ   Sequence [illegible] AA;

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches [illegible]80; Conservative [illegible]12; Mismatches [illegible]

Qy        67 [illegible]EVA[illegible]KAL[illegible]NN[illegible]WKK[illegible]
Db       231 [illegible]
Qy       127 [illegible]EMDREY[illegible]RME[illegible]WNPNVK[illegible]
Db       291 [illegible]
Qy       187 [illegible]KRR[illegible]AGM[illegible]EMP[illegible]VI[illegible]
Db       351 [illegible]
Qy       247 [illegible]WEPK[illegible]N[illegible]ANB[illegible]
Db       411 K[illegible]

RESULT 12
[illegible]726
ID   AAB4[illegible] standard; [illegible] AA.
XX
AC   AAB4[illegible];
XX
DT   08-FEB-[illegible]01 [illegible]
XX
DE   Human cancer [illegible] sequence [illegible]
XX
KW   Human; cancer [illegible]
KW   diagnosis; cytostatic; [illegible]
KW   antidiabetic; anti[illegible]
KW   antiinflammatory; [illegible]
KW   dermatological; [illegible]
KW   vasotropic; [illegible]
KW   immune disorder; [illegible]
KW   allergic reaction; [illegible]
KW   haemostatic; [illegible]
KW   neurological [illegible]
XX
OS   Homo sapiens.
XX
PN   WO200[illegible]
XX
PD   [illegible]
XX
```

```
[illegible]
XX
[illegible] MA[illegible]
XX
PA   [illegible]HUMA[illegible]
XX
[illegible]
XX
PT   [illegible]
PT   [illegible]
XX
PS   Claim [illegible]
XX
CC   AA[illegible]
CC   [illegible]
CC   the present invention.
XX
SQ   Sequence [illegible] AA;

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches [illegible]

Qy   [illegible]
Db   [illegible]
Qy   [illegible]
Db   [illegible]
Qy   [illegible]
Db   [illegible]

RESULT [illegible]
AA[illegible]
ID   AA[illegible] AA.
XX
AC   AA[illegible]
XX
DT   [illegible]
XX
DE   [illegible]
XX
KW   [illegible]
KW   [illegible]
XX
[illegible]
XX
```

```
PN   EP1033401-A2.
XX
PD   06-SEP-2000.
XX
PF   21-FEB-2000; 2000EP-0200619.
XX
PR   26-FEB-1999;    99US-0122497.
XX
PA   (GEST ) GENSET.
XX
PI   Dumas Milne Edwards J,  Duclert A,  [illegible] J.
XX
DR   WPI; 2000-500381/45.
DR   N-PSDB; AAC01736.
XX
PT   New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT   obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT   diagnostic, forensic, gene therapy and chromosome mapping procedures
XX
PS   Claim 13; SEQ ID 5811; 71pp + CD-ROM; English.
XX
CC   The present sequence is a polypeptide encoded by one of a large number
[illegible]   of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs
[illegible]   were prepared from total human RNAs or polyA+ RNAs derived from [illegible]
CC   different tissues. EST sequences usually correspond mainly to the 3'
CC   untranslated region (UTR) of the mRNA because they are often obtained
CC   from oligo-dT primed cDNA libraries. Such ESTs are not well suited for
CC   isolating cDNA sequences derived from the 5' ends of mRNAs and even in
CC   those cases where longer cDNA sequences have been obtained, the full 5'
CC   UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'
CC   ends and can therefore be used to obtain full length cDNAs and genomic
CC   DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and
CC   chromosome mapping procedures. They are used to obtain upstream
CC   regulatory sequences and to design expression and secretion vectors.
XX
SQ   Sequence   69 AA;


  Query Match             14.2%;  Score [illegible];  DB 21;  Length [illegible];
  Best Local Similarity   72.6%;  Pred. No. [illegible];
  Matches   45;  Conservative    4;  Mismatches   [illegible];  Indels   [illegible];  Gaps   [illegible];

Qy        1 MFLAIFKLCACSSYRHMPNMKGLRHQAVLA[illegible]NWRALCDSSP[illegible]WMGQVPRRS[illegible]SL[illegible]
              | |||||||||||||||||||||||||| |||:||  |  || ||||   :|     : ||      |:
Db        1 mllatfklcaqssyrhmrnmkglrqqa[illegible]alqqptps[illegible]xlnqv[illegible]ysv[illegible]

Qy       58 LG 59
            ||
Db       61 lq 62


[illegible]LT 14
[illegible]53294
ID   AAB53294 standard; Protein; 239 AA
XX
AC   AAB53294;
XX
DT   09-MAR-2001  (first entry)
XX
DE   Human colon cancer antigen protein sequence SEQ ID NO:[illegible].
XX
KW   Human; colon cancer; colon cancer antigen; diagnosis; detection;
KW   identification; cytostatic; cardioactive; neuroprotective; [illegible];
KW   immunomodulatory; muscular; gynaecological; gastrointestinal;
KW   nephrotropic; antiinfective; antibacterial; gene therapy; wound;
KW   neural disorder; immune system disorder; muscular disorder;
KW   reproductive disorder; gastrointestinal disorder; renal disorder;
KW   infectious disease; cardiovascular disorder.
XX
OS   Homo sapiens.
XX
PN   WO200055351-A1.
XX
```

```
PD   21-SEP-2000.
XX
PF   08-MAR-2000; 2000WO-US05883.
XX
PR   12-MAR-1999;    99US-0124270.
XX
PA   (HUMA ) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Ruben SM;
XX
DR   WPI; 2000-[illegible]/55.
DR   N-PSDB; AAC[illegible].
XX
PT   Colon cancer associated gene sequences, referred to as colon cancer
PT   antigens, useful for the treatment, prevention, and diagnosis of colon
PT   disorders such as colon cancer
XX
PS   Claim 11; Page 1385; 2104pp; English.
XX
CC   AAC97991 to AAC98765 encode the human colon cancer associated proteins,
CC   called human colon cancer antigens, given in AAB53234 to AAB54006. The
CC   human colon cancer antigens can have cytostatic, cardioactive, muscular,
CC   neuroprotective, immunomodulatory, gynaecological, gastrointestinal,
CC   vulnerary, nephrotropic, antiinfective and antibacterial activities, and
CC   can be used in gene therapy. The colon cancer antigen polynucleotides,
CC   proteins and antibodies to the proteins are useful for the prevention,
CC   treatment and diagnosis of colon disorders, such as colon cancer. The
CC   polynucleotides may be used in diagnostics and research, such as for
CC   chromosome identification, and as hybridisation probes. The proteins
CC   may also be used to prevent diseases such as neural disorders, immune
CC   system disorders, muscular disorders, reproductive disorders,
CC   gastrointestinal disorders, wounds, renal disorders, infectious
CC   diseases, and cardiovascular disorders. AAC98764 to AAC98772 and
CC   AAB54007 represent sequences used in the exemplification of the present
CC   invention.
XX
SQ   Sequence   239 AA;


  Query Match             13.[illegible]%;  Score [illegible];  DB 21;  Length 239;
  Best Local Similarity   26.3%;  Pred. No. [illegible];
  Matches   55;  Conservative   42;  Mismatches   [illegible];  Indels   1[illegible];  Gaps    7;

Qy       7[illegible] [illegible] 1[illegible]2
Db       [illegible] [illegible]lyrqealvyatleev 9[illegible]

Qy      1[illegible] YEELVDRMEAM[illegible] EWNPNVKE[illegible]KV[illegible] 1[illegible]
Db       91 wdev[illegible]kp[illegible]qirvkwdenvtq[illegible] 14[illegible]

Qy      189 [illegible]WILSIFI 246
Db      148 [illegible] 206

Qy      217 KGWLPK[illegible]EFANH[illegible]KKF[illegible] [illegible]
Db      207 sqylp[illegible]vdsfftprsmt[illegible]tyanl[illegible] 2[illegible]


RESULT  15
AAB43727
ID   AAB43727 standard; Protein; 1[illegible] AA
XX
AC   AAB43727;
XX
DT   08-FEB-2001  (first entry)
XX
DE   Human cancer associated protein sequence SEQ ID NO:1172.
XX
KW   Human; cancer associated gene; cancer antigen; detection; cancer;
KW   diagnosis; cytostatic; proliferative; vulnerary; immunomodulator;
KW   antidiabetic; antiasthmatic; antiinflammatory; antiarthritic; antiviral;
KW   antiinflammat[illegible]
KW   dermatological[illegible]
KW   vasotropic[illegible]
KW   immune disorder[illegible] haematopoiet[illegible]
KW   allergic [illegible]
KW   haemostatic[illegible]
KW   neur[illegible]
XX
[illegible]   Homo sapiens.
XX
PN   WO200[illegible]
XX
PD   21-SEP-2000.
XX
PF   08-MAR-200[illegible]   [illegible]W[illegible]
XX
PR   12-MAR-1999; [illegible]
[illegible]
     (HUMA-) HUMAN GENOME [illegible] INC.
XX
PI   Rosen CA, [illegible] SM
XX
DR   WPI; 2000-[illegible]
DR   N-PSDB; AA[illegible]
XX
PT   Novel isolated [illegible]
PT   useful for treating [illegible]
XX
[illegible]   Claim [illegible]; Page [illegible]; [illegible] pp; English.
XX
[illegible]   AA[illegible] to AA[illegible]
[illegible]   in AAB4[illegible] to AAB4[illegible]
[illegible]   tissues and [illegible]. The genes are expressed [illegible]
[illegible]   include cytostatic, [illegible]
[illegible]   antidiabetic, [illegible]
[illegible]   antiinflammat[illegible]
[illegible]   dermatological, [illegible]
[illegible]   nootropic; vasotropic, [illegible]
[illegible]   polynucleotides and polypeptides can be used [illegible]
[illegible]   amelio[illegible]
[illegible]   Polynucleotides, polypeptides, [illegible]
[illegible]   the present invention may be used [illegible]
[illegible]   or inhibiting the proliferation, [illegible]
[illegible]   immune cells [illegible]
[illegible]   disorders, [illegible]
[illegible]   rejection, [illegible]
[illegible]   inflammation, [illegible]
[illegible]   bacterial or viral infections. The peptide[illegible]
[illegible]   agonists and [illegible]
[illegible]   AA[illegible]784[illegible] and AAB[illegible]
[illegible]   the present invention.
XX
SQ   Sequence   [illegible] AA;

  Query Match   [illegible]
  Best Local Similarity   [illegible]
  Matches   [illegible]   Mismatches   [illegible]

Qy   [illegible] A[illegible]SSP[illegible] WMQ[illegible]
Db   [illegible]
Qy   [illegible] GEVAMQKA[illegible]NQ[illegible]
Db   [illegible]
Qy   [illegible] DEMFAMGEWNE[illegible]AAAAA[illegible]
Db   [illegible]
```

```
                       GenCore version 4.5
              Copyright (c) 1993 - 2001 Compugen Ltd.

OM protein - protein search, using sw model

Run on:          August 22, 2001, 14:30:57 ; Search time 35.8 Seconds
                                            (without alignments)
                                            10.525 Million cell updates/sec

Title:           US-08-659-254-8
Perfect score:   62
Sequence:        1 ILNNQEGWKKE 11

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters:       93435

 mum DB seq length: 0
 mum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       SwissProt_39:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                              SUMMARIES
                  %
Result          Query
  No.   Score   Match Length DB  ID                        Description
-----------------------------------------------------------------------------
   1      62  100.0    284  1  STAR_MOUSE                 P51557 mus musculu
   2      62  100.0    284  1  STAR_RAT                   P97826 rattus norv
   3      57   91.9    284  1  STAR_MESAU                 P79111 mesocricetu
   4      57   91.9    285  1  STAR_HUMAN                 P49675 homo sapien
   5      57   91.9    285  1  STAR_PIG                   Q28996 sus scrofa
   6      52   83.9    285  1  STAR_HORSE                 [illegible] equus cabal
   7      51   82.3    285  1  STAR_BOVIN                 Q28918 bos taurus
   8      43   69.4    157  1  UBC3_SCHPO                 P40984 schizosacch
   9      40   64.5    620  1  YG18_YEAST                 P53207 saccharomyc
  10      40   64.5   1511  1  PDR5_YEAST                 P33302 saccharomyc
 11      40   64.5   1529  1  PDRF_YEAST                 Q04182 saccharomyc
 12      39   62.9   1501  1  CDR3_CANAL                 O42690 candida alb
  13      39   62.9   1742  1  CHNA_CAL[illegible]        [illegible]
  14      38   61.3    451  1  MYBH_DIC                   P34127 dictyosteli
  15      38   61.3    475  1  TBE_HUMAN                  Q9ujt0 homo sapien
  16      38   61.3   1588  1  ARO1_YEAST                 P08566 s pentatune
  17      37   59.7    322  1  GRP2_MOUSE                 [illegible]89100 m grb2-rela
  18      37   59.7    358  1  25A1_RAT                   Q05961 rattus norv
  19      37   59.7    358  1  OAS1_RAT                   Q05961 rattus norv
  20      37   59.7    800  1  GUN_BACS                   P06564 bacillus sp
  21      37   59.7   1271  1  Y338_MYCPN                 P47580 mycoplasma
  22      36   58.1    112  1  PSBW_CYAPA                 P48264 cyanophora
  23      36   58.1    194  1  RK24_PEA                   P11893 pisum sati
  24      36   58.1    330  1  GRP2_HUMAN                 [illegible]75791 h grb2 rela
  25      36   58.1    353  1  YFH7_YEAST                 P43591 saccharomyc
  26      36   58.1    368  1  LYS1_SCHPO                 Q09694 schizosacch
  27      36   58.1    458  1  NUON_RICPR                 Q9zd13 rickettsia
  28      36   58.1    505  1  TCMO_CATRO                 P48522 catharanthu
  29      36   58.1    799  1  ZY11_CAEEL                 P21541 caenorhabdi
  30      36   58.1    865  1  CYAA_PROMI                 Q59685 proteus mir
  31      36   58.1    877  1  YBGG_ECOLI                 P54746 escherichia
  32      36   58.1   3587  1  SRF1_BACSU                 P27206 bacillus su
  33      36   58.1   3587  1  SRF2_BACSU                 Q04747 bacillus su
```

```
  34      35   56.5    269  1  GEM2_RAT                   Q9qzp1 rattus norv
  35      35   56.5    280  1  GEM2_HUMAN                 O14893 homo sapien
  36      35   56.5    331  1  CG12_HUMAN                 O95236 homo sapien
  37      35   56.5    393  1  DOM3_CAEEL                 Q10660 caenorhabdi
  38      35   56.5    406  1  TRKA_METMA                 P39449 methanosarc
  39      35   56.5    417  1  PRI_MOUSE                  P20664 mus musculu
  40      35   56.5    420  1  PRI_HUMAN                  P49642 homo sapien
  41      35   56.5    439  1  CKI[illegible]             [illegible]74135 schizosacch
  42      35   56.5    446  1  CKI[illegible]             P40233 schizosacch
  43      35   56.5    489  1  PEN3_ADEM1                 O10439 mouse adeno
  44      35   56.5    489  1  PT94_YEAST                 P07390 saccharomyc
  45      35   56.5    563  1  HNM1_YEAST                 P19807 saccharomyc

                              ALIGNMENTS

RESULT    1
STAR_MOUSE
ID   STAR_MOUSE     STANDARD;      PRT;   284 AA.
AC   P51557;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN PRECURSOR (LUTEINIZING HORMONE-
DE   INDUCED PROTEIN).
GN   STAR.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID 10090;
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX   MEDLINE 95058616, PubMed 7961770;
RA   Clark B.J., Wells J., King S.R., Stocco D.M.;
RT   "The purification, cloning, and expression of a novel luteinizing
RT   hormone-induced mitochondrial protein in MA-10 mouse Leydig tumor
RT   cells. Characterization of the steroidogenic acute regulatory protein
RT   (STAR).";
RL   J. Biol. Chem. 269:28314-28322(1994).
CC   -!- FUNCTION: PLAYS A KEY ROLE IN STEROID HORMONE SYNTHESIS BY
CC       ENHANCING THE METABOLISM OF CHOLESTEROL INTO PREGNENOLONE.
CC       TRANSLOKEER THAT BINDS TO AND TRANSPORT CHOLESTEROL THROUGH THE
CC       INTERMEMBRANE SPACE OF THE MITOCHONDRION (PROBABLE).
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL.
CC   -!- SIMILARITY: BELONGS TO THE STAR FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See [illegible]
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; L36062; AAB94783.1; -.
DR   MGD; MGI:102760; Star.
DR   InterPro; IPR000799; -.
DR   InterPro; IPR002913; -.
DR   Pfam; PF01852; START; 1.
DR   PRINTS; PR00978; STARPROTEIN.
KW   Lipid binding; Lipid transport; Transport; Steroidogenesis;
KW   Mitochondrion; Transit peptide.
FT   TRANSIT       1      ?       MITOCHONDRION (POTENTIAL).
FT   CHAIN         ?    284       STEROIDOGENIC ACUTE REGULATORY PROTEIN
FT   DOMAIN      170    174       POLY-ALA.
SQ   SEQUENCE   284 AA;  31626 MW;  9BB13F52E4439B72 CRC64;

  Query Match             100.0%;  Score 62;  DB 1;  Length 284;
  Best Local Similarity   100.0%;  Pred. No. [illegible]60095;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy        1 [illegible]NNQE[illegible]WKK[illegible] 11
            ||||||||||
Db       88 [illegible]NNQE[illegible]WKK[illegible] [illegible]


RESULT  2
STAR_RAT
ID   STAR_RAT       STANDARD;      PRT;   [illegible] AA.
AC   P97826; P97[illegible];
DT   01 NOV 1997 (Rel. 35, Created)
DT   01 NOV 1997 (Rel. 35, Last sequence update)
DT   30 MAY 2000 (Rel. 39, Last annotation update)
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN [illegible]
GN   STAR.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; [illegible]
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID [illegible]
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN WISTAR; TISSUE [illegible];
RX   MEDLINE 9746[illegible]; PubMed [illegible];
RA   Mizutani T., Sakata [illegible], Minegishi [illegible], Wakatsuki [illegible] [illegible]
RT   "Molecular cloning, characterization and [illegible]
RT   steroidogenic acute regulatory protein (StAR) [illegible]"
RL   Life Sci. 61:14[illegible]-15[illegible](1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN SPRAGUE-DAWLEY; TISSUE [illegible];
RX   MEDLINE 9716[illegible]; PubMed [illegible];
RA   Lee H.K., Ahn R.[illegible], Kwon H.B., [illegible]
RT   "Nucleotide sequence of rat steroidogenic acute regulatory protein
RT   complementary DNA.";
RL   Biochem. Biophys. Res. Commun. [illegible](1997).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN SPRAGUE-DAWLEY; TISSUE Adrenal gland;
RX   MEDLINE 981848[illegible]; PubMed [illegible];
RA   Ariyoshi N., Kim Y.[illegible], Artemenko I., Bhattacharyya K.K.,
RA   Jefcoate C.R.
RT   "Characterization of the rat Star gene that encodes the [illegible]
RT   3.5 kilobase pair mRNA. [illegible]
RT   precedes elevation of Star mRNA and protein."
RL   J. Biol. Chem. [illegible](1998).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN SPRAGUE-DAWLEY; TISSUE Adrenal [illegible];
RX   MEDLINE 99048[illegible]; PubMed [illegible];
RA   Furukawa A., Miyatake A., [illegible]
RT   "Steroidogenic acute regulatory protein (StAR) [illegible]
RT   constitutively expressed in the adult rat [illegible]
RT   colocalization of StAR, cytochrome P-450scc (CYPXIA[illegible]) [illegible]
RT   hydroxysteroid dehydrogenase in the rat brain."
RL   J. Neurochem. 71:2231-2238(1998).
CC   -!- FUNCTION: PLAYS A KEY ROLE IN STEROID HORMONE SYNTHESIS BY
CC       ENHANCING THE METABOLISM OF CHOLESTEROL INTO PREGNENOLONE.
CC       TRANSPORTER THAT BINDS [illegible] AND TRANSFERS [illegible]
CC       INTERMEMBRANE SPACE OF THE MITOCHONDRION (PROBABLE).
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL.
CC   -!- SIMILARITY: BELONGS TO THE STAR FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.[illegible]
CC   or send an email to license@[illegible]).
CC   --------------------------------------------------------------------------
DR   EMBL; AB[illegible]; BAA19[illegible]; -.
DR   EMBL; U76419; AAB[illegible]; -.
DR   EMBL; AF[illegible]; AA[illegible]; -.
DR   EMBL; AB[illegible]; BAA3[illegible]; -.
```

```
DR   [illegible]
DR   [illegible]
DR   [illegible]
DR   PRINTS; PR[illegible]; STARPROTEIN.
KW   Lipid-binding; Lipid transport; Transport; Steroidogenesis;
KW   Mitochondrion; Transit peptide.
FT   TRANSIT      [illegible]       MITOCHONDRION (POTENTIAL).
FT   CHAIN        [illegible]  284   STEROIDOGENIC ACUTE REGULATORY PROTEIN.
FT   CONFLICT     [illegible]       [illegible]
FT                                 [illegible]
FT                                 [illegible]
SQ   SEQUENCE   284 AA;  [illegible] MW;  [illegible] CRC64;

  Query Match           [illegible]  Score [illegible];  Length 284;
  Best Local Similarity  [illegible];  Pred. No. [illegible];
  Matches  [illegible];  Conservative  [illegible];  Mismatches  [illegible];  Indels  [illegible];  Gaps  [illegible];

Qy        1 [illegible]NNQ[illegible]WKK[illegible] [illegible]

Db       88 [illegible]NNQ[illegible]WKK[illegible] [illegible]


RESULT  3
STAR_MESAU
ID   STAR_MESAU     STANDARD;      PRT;   [illegible] AA.
AC   [illegible];
DT   01 NOV 1997 (Rel. 35, Created)
DT   01 NOV 1997 (Rel. 35, Last sequence update)
DT   30 MAY [illegible] (Rel. 39, Last annotation update)
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN PRECURSOR.
GN   STAR.
OS   Mesocricetus auratus (Golden hamster).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC   Mesocricetus.
OX   NCBI_TaxID [illegible];
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE [illegible]; PubMed [illegible];
RA   [illegible] A., [illegible] M., [illegible], Lebeux [illegible]
RA   [illegible]
RT   "Adrenocorticotropin [illegible] regulates the levels of the steroidogenic acute
RT   regulatory (STAR) protein mRNA in hamster adrenals.";
RL   Endocr. [illegible]
CC   -!- FUNCTION: PLAYS A KEY ROLE IN STEROID HORMONE SYNTHESIS BY
CC       ENHANCING THE METABOLISM OF CHOLESTEROL INTO PREGNENOLONE.
CC       TRANSPORTER THAT BINDS [illegible] AND TRANSFERS [illegible] THE [illegible]
CC       INTERMEMBRANE SPACE OF THE MITOCHONDRION (PROBABLE).
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL.
CC   -!- SIMILARITY: BELONGS TO THE STAR FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.[illegible]/announce/
CC   or send an email to license@[illegible]).
CC   --------------------------------------------------------------------------
DR   EMBL; [illegible]; AA[illegible]; -.
DR   [illegible]
DR   [illegible]
DR   [illegible]
DR   PRINTS; PR[illegible]; STARPROTEIN.
KW   Lipid-binding; Lipid transport; Transport; Steroidogenesis;
KW   Mitochondrion; Transit peptide.
FT   TRANSIT      [illegible]       MITOCHONDRION (POTENTIAL).
FT   CHAIN        [illegible]  284   STEROIDOGENIC ACUTE REGULATORY PROTEIN.
SQ   SEQUENCE   284 AA;  [illegible] MW;  [illegible] CRC64;


  Query Match           [illegible];  Score [illegible];  DB [illegible];  Length 284;
  Best Local Similarity   90.9%;  Pred. No. [illegible];
  Matches   10;  Conservative    1;  Mismatches    0;  Indels   [illegible];  Gaps   [illegible]

Qy       1 ILNNQEGWKKE 11
           ||:||||||||
Db      88 ILSNQEGWKKE 98


RESULT   4
STAR_HUMAN
ID   STAR_HUMAN     STANDARD;      PRT;   28[illegible] AA
AC   P49675; Q16396;
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN PRECURSOR.
GN   STAR.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
     SEQUENCE FROM N.A.
     TISSUE=Adrenal cortex;
RX   MEDLINE 95281540; PubMed 7761400;
RA   Sugawara T., Holt J.A., Driscoll D., Strauss J.F. III, Lin D.,
RA   Miller W.L., Patterson D., Clancy K.P., Hart I.M., Clark B.J.,
RA   Stocco D.M.;
RT   "Human steroidogenic acute regulatory protein: functional activity in
RT   COS-1 cells, tissue-specific expression, and mapping of the
RT   structural gene to 8p11.2 and a pseudogene to chromosome 13.";
RL   Proc. Natl. Acad. Sci. U.S.A. 92:4778-4782(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE-Placenta;
RX   MEDLINE 96038208; PubMed 7547998;
RA   Sugawara T., Lin D., Holt J.A., Martin K.O., Javitt N.B., Miller W.L.,
RA   Strauss J.F. III;
RT   "Structure of the human steroidogenic acute regulatory protein (StAR)
RT   gene: StAR stimulates mitochondrial cholesterol 27-hydroxylase
RT   activity.";
RL   Biochemistry 34:12506-12512(1995).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96008588; PubMed=7548191;
RA   Gradi A., Tang-Wai R., McBride H.M., [illegible] L.L., Shore G.C.,
RA   Pelletier J.;
RT   "The human steroidogenic acute regulatory (StAR) gene is expressed in
RT   the urogenital system and encodes a mitochondrial polypeptide.";
RL   Biochim. Biophys. Acta 1259:228-233(1995).
     [4]
     SEQUENCE FROM N.A.
     TISSUE-Brain;
RA   Yu W., Sarginson J., Gibbs R.A.;
RL   Submitted (NOV-1997) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   VARIANT CLAH LEU-182
RX   MEDLINE=97096837; PubMed 8948562.
RA   Bose H.S., Sugawara T., Strauss J.F. III, Miller W.L.;
RT   "The pathophysiology and genetics of congenital lipoid adrenal
RT   hyperplasia.";
RL   New Engl. J. Med. 335:1870-1878(1996).
RN   [6]
RP   VARIANT ALA-203.
RX   MEDLINE-98112495; PubMed 9452116.
RA   Katsumata N., Tanae A., Shinagawa T., Nagashima-Miyokawa A.,
RA   Shimizu M., Yasunaga T., Tanaka T., Hibi I.;
RT   "A novel frameshift mutation 840del[illegible] and a novel polymorphism [illegible] in
RT   the steroidogenic acute regulatory protein gene in a Japanese patient
RT   with congenital lipoid adrenal hyperplasia.";
RL   Hum. Mutat. Suppl. 1:S304-S307(1998).
CC   -!- FUNCTION: PLAYS A KEY ROLE IN STEROID HORMONE SYNTHESIS BY
CC       ENHANCING THE METABOLISM OF CHOLESTEROL INTO PREGNENOLONE.
CC       TRANSPORTER THAT BINDS TO AND TRANSFER CHOLESTEROL THROUGH THE
CC       INTERMEMBRANE SPACE OF THE MITOCHONDRION (PROBABLE).
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN GONADS, ADRENAL CORTEX AND
CC       KIDNEY.
CC   -!- DISEASE: DEFECTS IN STAR CAUSE CONGENITAL LIPOID ADRENAL
CC       HYPERPLASIA (CLAH), THE MOST SEVERE FORM OF ADRENAL HYPERPLASIA.
CC   -!- SIMILARITY: BELONGS TO THE STAR FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement [illegible]
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; U17280; AAC50141.1; -.
DR   EMBL; U29105; AAC50234.1; -.
DR   EMBL; U29099; AAC50234.1; JOINED.
DR   EMBL; U29100; AAC50234.1; JOINED.
DR   EMBL; U29101; AAC50234.1; JOINED.
DR   EMBL; U29102; AAC50234.1; JOINED.
DR   EMBL; U29103; AAC50234.1; JOINED.
DR   EMBL; U29104; AAC50234.1, JOINED.
DR   EMBL; S79669; AAB35726.1; -.
DR   EMBL; AF035277; AAB88174.1; -.
DR   MIM; 600617; -.
DR   MIM; 201710; -.
DR   InterPro; IPR000799; -.
DR   InterPro; IPR002913; -.
DR   Pfam; PF01852; START; 1.
DR   PRINTS; PR00978; STARPROTEIN.
KW   Lipid-binding; Lipid transport; Transport; Steroidogenesis;
KW   Mitochondrion; Transit peptide; Disease mutation; Polymorphism.
FT   TRANSIT       1      ?       MITOCHONDRION (POTENTIAL).
FT   CHAIN         ?    285       STEROIDOGENIC ACUTE REGULATORY PROTEIN.
FT   VARIANT     182    182       R -> L (IN CLAH).
FT                                /FTId=VAR_005627.
FT   VARIANT     203    203       D -> A.
FT                                /FTId=VAR_005628.
SQ   SEQUENCE   285 AA;  31958 MW;  [illegible]38AD [illegible]64;


  Query Match             91.9%;  Score [illegible];  DB [illegible];  Length 285;
  Best Local Similarity   90.9%;  Pred. No. [illegible]71;
  Matches   10;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 ILNNQEGWKKE 11
           ||:||||||||
Db      89 ILSNQEGWKKE 99


RESULT   5
STAR_PIG
ID   STAR_PIG       STANDARD;      PRT;   28[illegible] AA.
AC   Q28996; Q95259;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN PRECURSOR.
GN   STAR.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Ovary;
RX   MEDLINE=97200784; PubMed=9048613.
RA   Pilon N., Daneau I., Brisson C., Ethier J.F., Lussier J.G.,
RA   Silversides D.W.;
RT   "Porcine and bovine steroidogenic acute regulatory protein (StAR)
```

```
RL   gene expression during gestation."
RL   Endocrinology [illegible]
RN   [2]
RP   SEQUENCE OF [illegible] FROM N.A.
RC   TISSUE=Corpus luteum;
RX   MEDLINE=971[illegible]; PubMed=[illegible];
RA   Balasubramanian K., [illegible] H.A., [illegible]
RA   Veldhuis J.D.;
RT   "Regulation of porcine granulosa cell steroidogenic acute regulatory
RT   protein (StAR) by insulin-like growth factor [illegible]
RT   follicle-stimulating hormone or protein kinase A [illegible]"
RL   Endocrinology 138:4[illegible](1997).
CC   -!- FUNCTION: PLAYS A KEY ROLE IN STEROID HORMONE SYNTHESIS BY
CC       ENHANCING THE METABOLISM OF CHOLESTEROL INTO PREGNENOLONE.
CC       TRANSPORTER THAT BINDS [illegible] AND TRANSFER [illegible]
CC       INTERMEMBRANE SPACE OF THE MITOCHONDRION (PROBABLE).
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN ADULT AN[illegible]
CC       TISSUES, INCLUDING ADULT TESTES AND OVARIES AND ADULT ADRENAL
CC       GLANDS AS WELL AS STEROIDOGENIC [illegible] PREGNANCY [illegible]
CC       DEVELOPING FETAL TESTES, CORPUS LUTEUM AND PREGNANCY [illegible]
CC       THE FETAL OVARY.
CC   -!- SIMILARITY: BELONGS TO THE STAR FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. [illegible]
CC   between the Swiss Institute of Bioinformatics [illegible] EMBL [illegible]
CC   the European Bioinformatics Institute. [illegible]
CC   use by non-profit institutions [illegible]
CC   modified and this statement is not removed. [illegible]
CC   entities requires a license agreement (See [illegible]
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; [illegible]; AAB[illegible].
DR   EMBL; [illegible]195; AAB[illegible].
DR   InterPro; IPR[illegible];
DR   InterPro; IPR00[illegible];
DR   Pfam; PF018[illegible]; START; [illegible]
DR   PRINTS; PR00[illegible]; STARPROTEIN.
KW   Lipid-binding; Lipid transport; Transport; Steroidogenesis;
KW   Mitochondrion; Transit peptide.
FT   TRANSIT       [illegible]       MITOCHONDRION (POTENTIAL).
FT   CHAIN         [illegible]       STEROIDOGENIC ACUTE REGULATORY PROTEIN.
FT   CONFLICT      [illegible]
SQ   SEQUENCE   28[illegible] AA;  [illegible] MW;  [illegible] CRC64;

  Query Match         [illegible]
  Best Local Similarity [illegible]  Pred. No. [illegible]
  Matches [illegible]  Conservative [illegible]  Mismatches [illegible]  Indels [illegible]  Gaps [illegible]

Qy          1 IINNQEGWKKE [illegible]
              [illegible]
Db         80 IISNQEGWKKE [illegible]


RESULT 6
STAR_HORSE
ID   STAR_HORSE     STANDARD;      PRT;   [illegible] AA.
AC   O46689;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN [illegible]
GN   STAR.
OS   Equus caballus (Horse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; [illegible]
OC   Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX   NCBI_TaxID=9796;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Follicle;
RX   MEDLINE=9[illegible]2448; PubMed=[illegible];
RA   Kerban A., Boerboom [illegible]
RL   "Human [illegible]
RL   steroidogenic acute regulatory protein [illegible]
RL   theca interna of [illegible] of equine preovulatory follicles.";
RL   Endocrinology [illegible]
CC   -!- FUNCTION: PLAYS A KEY ROLE IN STEROID HORMONE SYNTHESIS BY
CC       ENHANCING THE METABOLISM OF CHOLESTEROL INTO PREGNENOLONE.
CC       TRANSPORTER THAT BINDS [illegible] AND TRANSFER [illegible] THROUGH THE
CC       INTERMEMBRANE SPACE OF THE MITOCHONDRION (PROBABLE).
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL.
CC   -!- SIMILARITY: BELONGS TO THE STAR FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. [illegible]
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AF[illegible]; AAC[illegible].
DR   EMBL; AF[illegible]; AAC[illegible].
DR   InterPro; IPR[illegible];
DR   InterPro; IPR[illegible];
DR   Pfam; PF[illegible]; START; [illegible]
DR   PRINTS; PR00[illegible]; STARPROTEIN.
KW   Lipid-binding; Lipid transport; Transport; Steroidogenesis;
KW   Mitochondrion; Transit peptide.
FT   TRANSIT       [illegible]       MITOCHONDRION (POTENTIAL).
FT   CHAIN         [illegible]  285  STEROIDOGENIC ACUTE REGULATORY PROTEIN.
SQ   SEQUENCE   285 AA;  [illegible] MW;  [illegible] CRC64;

  Query Match         [illegible]  Score [illegible]  Length 285;
  Best Local Similarity [illegible]  Pred. No. [illegible]
  Matches [illegible]  Conservative [illegible]  Mismatches [illegible]  Indels [illegible]  Gaps [illegible]

Qy          [illegible] INNQEGWKKE

Db         [illegible] INNQEGWKKE [illegible]


RESULT 7
STAR_BOVIN
ID   STAR_BOVIN     STANDARD;      PRT;   [illegible] AA.
AC   [illegible];
DT   [illegible] NOV [illegible] (Rel. [illegible], Created)
DT   15-DEC-[illegible] (Rel. [illegible], Last sequence update)
DT   [illegible] MAY [illegible] (Rel. [illegible], Last annotation update)
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN PRECURSOR.
GN   STAR.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Corpus luteum;
RX   MEDLINE=[illegible]; PubMed=[illegible];
RA   Hartung S., Rust W., Balvers M., Ivell R.;
RT   "Molecular cloning and in vivo expression of the bovine steroidogenic
RT   acute regulatory protein [illegible]."
RL   Biochem. Biophys. Res. Commun. 215:646-[illegible](1995).
RN   [2]
RP   REVISIONS [illegible] AND [illegible]
RA   Hartung S., Rust W., Balvers M., Ivell R.;
RL   Submitted (MAY-[illegible]) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: PLAYS A KEY ROLE IN STEROID HORMONE SYNTHESIS BY
CC       ENHANCING THE METABOLISM OF CHOLESTEROL INTO PREGNENOLONE.
CC       TRANSPORTER THAT BINDS [illegible] AND TRANSFER [illegible] THROUGH THE
CC       INTERMEMBRANE SPACE OF THE MITOCHONDRION (PROBABLE).
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL.
CC   -!- TISSUE SPECIFICITY: CORPUS LUTEUM AND ADRENAL GLAND.
CC   -!- SIMILARITY: BELONGS TO THE STAR FAMILY
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; Y17259; CAA76717.1; -.
DR   EMBL; Y17260; CAA76718.1; -.
DR   InterPro; IPR000799; -.
DR   InterPro; IPR002913; -.
DR   Pfam; PF01852; START; 1.
DR   PRINTS; PR00978; STARPROTEIN.
KW   Lipid-binding; Lipid transport; Transport; Steroidogenesis;
KW   Mitochondrion; Transit peptide.
FT   TRANSIT       1      ?       MITOCHONDRION (POTENTIAL).
FT   CHAIN         ?    285       STEROIDOGENIC ACUTE REGULATORY PROTEIN.
FT   CONFLICT     75     75       H -> Y (IN GENOMIC SEQUENCE).
FT   CONFLICT     99     99       E -> K (IN GENOMIC SEQUENCE).
[illegible]   SEQUENCE   285 AA;  31845 MW;  EAFD5C [illegible] 95E235 CRC64;

  Query Match             82.3%;  Score 51;  DB 1;  Length 285;
  Best Local Similarity   81.8%;  Pred. No. [illegible].078;
  Matches    9;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 ILNNQEGWKKE 11
              || :|||||||
Db         89 ILKDQEGWKKE 99


RESULT   8
UBC3_SCHPO
ID   UBC3_SCHPO     STANDARD;      PRT;   157 AA.
AC   P40984;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   UBIQUITIN-CONJUGATING ENZYME E2-18 KDA (EC 6.3.2.19)
DE   (UBIQUITIN-PROTEIN LIGASE HUS5) (UBIQUITIN CARRIER PROTEIN HUS5).
GN   HUS5 OR SPAC30D11.13.
OS   Schizosaccharomyces pombe (Fission yeast).
OC   Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC   Schizosaccharomycetales; Schizosaccharomycetaceae;
OC   Schizosaccharomyces.
OX   NCBI_TaxID=4896;
RN   [1]
[illegible]   SEQUENCE FROM N.A.
[illegible]   MEDLINE=95286695; PubMed=7768995;
RA   Al-Khodairy F., Enoch T., Hagman I.M., Carr A.M.;
RT   "The Schizosaccharomyces pombe hus5 gene encodes a ubiquitin
RT   conjugating enzyme required for normal mitosis.";
RL   J. Cell Sci. 108:475-486(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=972.
RA   Pearson D., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;
RL   Submitted (NOV-1995) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO
CC       OTHER PROTEINS. REQUIRED FOR EFFICIENT RECOVERY FROM DNA DAMAGE OR
CC       S-PHASE ARREST AND NORMAL MITOSIS. THIS MAY BE AS PART OF A
CC       CHECKPOINT INDEPENDENT RECOVERY PROCESS.
CC   -!- CATALYTIC ACTIVITY: ATP + UBIQUITIN + PROTEIN LYSINE = AMP +
CC       PYROPHOSPHATE + PROTEIN N-UBIQUITYLLYSINE.
CC   -!- PATHWAY: SECOND STEP IN UBIQUITIN CONJUGATION.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR
CC       UBIQUITIN-THIOLESTER FORMATION.
CC   -!- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY
CC       STRONGEST, TO YEAST UBC9.
```

```
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; X81846; CAA57438.1; -.
DR   EMBL; Z67961; CAA91899.1; -.
DR   HSSP; P50550; 1A3S.
DR   InterPro; IPR000608; -.
DR   Pfam; PF00179; UQ_con; 1.
DR   PROSITE; PS00183; UBIQUITIN_CONJUGAT_1; 1.
DR   PROSITE; PS50127; UBIQUITIN_CONJUGAT_2; 1.
KW   Ubiquitin conjugation; Ligase; DNA repair; Nuclear protein;
KW   Multigene family.
FT   BINDING      93     93       UBIQUITIN (BY SIMILARITY).
SQ   SEQUENCE   157 AA;  17989 MW;  4BF [illegible] 27F CRC64;

  Query Match             69.4%;  Score 4[illegible];  DB 1;  Length 157;
  Best Local Similarity   77.8%;  Pred. No. [illegible];
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 ILNNQEGWK 9
              ||| :||||
Db         96 ILNEEEGWK 104


RESULT   9
YG18_YEAST
ID   YG18_YEAST     STANDARD;      PRT;   620 AA.
AC   P53207;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   HYPOTHETICAL 71.4 KDA PROTEIN IN [illegible] INTERGENIC REGION.
GN   YGR013W.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288C;
RX   MEDLINE=97435481; PubMed=9290212;
RA   Rieger M., Brueckner M., Schaefer M., Mueller-Auer S.;
RT   "Sequence analysis of 203 kilobases from Saccharomyces cerevisiae
RT   chromosome VII.";
RL   Yeast 13:1077-1090(1997).
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; Z72798; CAA96996.1; -.
DR   SGD; S0003245; SNU71.
KW   Hypothetical protein.
SQ   SEQUENCE   620 AA;  71382 MW;  [illegible] CRC64;

  Query Match             64.5%;  Score 4[illegible];  DB 1;  Length 620;
  Best Local Similarity   60.0%;  Pred. No. 1[illegible];
  Matches    6;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 LNNQEGWKKE 11
              |::||||| :
```

```
Db      [illegible] QE [illegible] WK [illegible]


RESULT  16
PDR5_YEAST
ID   PDR5_YEAST     STANDARD;     PRT;   [illegible] AA.
AC   P33302;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   SUPPRESSOR OF [illegible]
GN   PDR5 OR STS1 OR YDR [illegible] OR LEM1 OR Y [illegible] W.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; [illegible]
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID [illegible];
RN   [1]
     SEQUENCE FROM N.A.
     STRAIN AB320;
RX   MEDLINE 94 [illegible]; PubMed [illegible]
RA   Bissinger P.H., Kuchler K.;
RT   "Molecular cloning and expression of the [illegible]
RT   STS1 gene product. A yeast ABC [illegible]
RT   resistance.";
RL   J. Biol. Chem. [illegible] 4180 [illegible]
RN   [2]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE
RC   STRAIN S288C
RX   MEDLINE 9412 [illegible]; PubMed [illegible]
RA   Balzi E., Wang M., Leterme S., [illegible]
RT   "PDR5, a novel yeast [illegible]
RT   controlled by the transcription [illegible]
RL   J. Biol. Chem. [illegible]
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN AH22;
RX   MEDLINE 951 [illegible]; PubMed [illegible]
RA   Hirata D., Yano K., [illegible]
RT   "Saccharomyces cerevisiae YDR1, which encodes a member of the A
RT   binding cassette (ABC) [illegible]
RT   resistance.";
RL   Curr. Genet. [illegible]
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN S288C / FY1679;
     Ayadi A., Bordonne R., [illegible] A., [illegible] A.,
     Tarassov I.A., [illegible]
RL   Submitted (MAY-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: TRANSPORTER [illegible]
CC       ESTABLISHED. CONFERS RESISTANCE [illegible]
CC       AND SULFOMETURON METHYL. EXHIBITS [illegible]
CC       ACTIVITY.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC   -!- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT [illegible]
CC       (ABC TRANSPORTERS) [illegible] PDR5 SUBFAMILY.
CC   -!- CAUTION: REF. [illegible] SEQUENCE DIFFERS FROM [illegible]
CC       FRAMESHIFT IN POSITION [illegible]
CC   --
CC   This SWISS-PROT entry is copyright. [illegible]
CC   between the Swiss Institute of Bioinformatics and the EMBL [illegible]
CC   the European Bioinformatics Institute. [illegible]
CC   use by non-profit [illegible]
CC   modified and this statement is not removed. [illegible]
CC   entities requires a license agreement [illegible]
CC   or send an email to license@isb-sib.ch).
CC   --
DR   EMBL; X74113; CAA52216.1; -.
DR   EMBL; L19922; AAB [illegible]
DR   EMBL; D26148; BAA [illegible] ALT_IN
DR   EMBL; U55020; AAC [illegible]
DR   EMBL; Z [illegible]; CAA [illegible]
DR   PIR; S34702; S341 [illegible]
DR   PIR; A4 [illegible]; A4 [illegible]
```

```
DR   [illegible]
DR   [illegible]
DR   [illegible]
DR   [illegible] ABC TRANSPORTER [illegible]
KW   ATP-binding; Transmembrane; [illegible]
FT   DOMAIN        [illegible]    CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM      [illegible]    POTENTIAL.
FT   TRANSMEM      [illegible]    POTENTIAL.
FT   TRANSMEM      [illegible]    POTENTIAL.
FT   TRANSMEM      [illegible]    POTENTIAL.
FT   TRANSMEM      [illegible]    POTENTIAL.
FT   TRANSMEM      [illegible]    POTENTIAL.
FT   DOMAIN        [illegible]    CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM      [illegible]    POTENTIAL.
FT   TRANSMEM      [illegible]    POTENTIAL.
FT   TRANSMEM      [illegible]    POTENTIAL.
FT   TRANSMEM      [illegible]    POTENTIAL.
FT   TRANSMEM      [illegible]    POTENTIAL.
FT   TRANSMEM      [illegible]    POTENTIAL.
FT   DOMAIN        [illegible]    CYTOPLASMIC (POTENTIAL).
FT   NP_BIND       [illegible]    ATP (POTENTIAL).
FT   DOMAIN        [illegible]    [illegible]
FT   CONFLICT      [illegible]    N [illegible] (IN REF. [illegible]
FT   CONFLICT      [illegible]    [illegible] (IN REF. [illegible]
FT   CONFLICT      [illegible]    [illegible] (IN REF. [illegible]
FT   CONFLICT      [illegible]    V (IN REF. [illegible]
FT   CONFLICT      [illegible]    MISSING (IN REF. [illegible]
FT   CONFLICT      [illegible]    H (IN REF. [illegible]
FT   CONFLICT      [illegible]    MISSING (IN REF. [illegible]
FT   CONFLICT      [illegible]    H (IN REF. [illegible]
SQ   SEQUENCE   [illegible] AA;  [illegible] MW;  [illegible]


  Query Match  [illegible]
  Best Local Similarity  [illegible]  Pred. No.  [illegible]
  Matches  [illegible]  Mismatches  [illegible]  Indels  [illegible]  Gaps  [illegible]

Qy         [illegible] NNL [illegible] WKK

Db         [illegible] NNV [illegible] WVK [illegible]


RESULT  [illegible]
PDR1_YEAST
ID   [illegible]_YEAST     STANDARD;     PRT;   [illegible] AA.
AC   [illegible]
DT   [illegible]-JUL-[illegible] (Rel. [illegible], Created)
DT   [illegible] (Rel. [illegible], Last sequence update)
DT   [illegible]-JUL-[illegible] (Rel. [illegible], Last annotation update)
DE   [illegible] DEPENDENT [illegible]
GN   [illegible] OR Y [illegible] W [illegible]
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; [illegible] Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID [illegible];
RN   [1]
RP   SEQUENCE FROM N.A.
RA   [illegible] M., [illegible] A., [illegible] A., [illegible] A., [illegible] A.,
RA   Carpenter [illegible] Chen [illegible] Chen [illegible] M., Chang [illegible] Duncan M.,
RA   [illegible] K., [illegible]
RA   Mc [illegible] Nakahara K., Nimarth A., [illegible] F., [illegible] R.X.,
RA   Roberts [illegible] Schramm S., Schroeder M., [illegible] Sherlett N.,
RA   Wilkinson A., [illegible] M., [illegible] Davis R.W.;
RL   Submitted (APR-[illegible]) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC   -!- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
CC       (ABC TRANSPORTERS) [illegible] SUBFAMILY.
CC   --
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use [illegible] was
CC   modified and this statement is not removed. [illegible]
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch)
CC   --------------------------------------------------------------------------
DR   EMBL; U32274; AAB64846.1; -.
DR   SGD; S0002814; PDR15.
DR   InterPro; IPR001617; -.
DR   Pfam; PF00005; ABC_tran; 2.
DR   PROSITE; PS00211; ABC_TRANSPORTER; 1.
KW   ATP-binding; Transmembrane; Glycoprotein; Transport.
FT   DOMAIN        1    531       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    532    552       POTENTIAL.
FT   TRANSMEM    568    588       POTENTIAL.
FT   TRANSMEM    618    638       POTENTIAL.
FT   TRANSMEM    643    663       POTENTIAL.
FT   TRANSMEM    700    720       POTENTIAL.
FT   TRANSMEM    784    804       POTENTIAL.
FT   DOMAIN      805   1219       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   1220   1240       POTENTIAL.
FT   TRANSMEM   1313   1333       POTENTIAL.
FT   TRANSMEM   1341   1361       POTENTIAL.
FT   TRANSMEM   1369   1389       POTENTIAL.
FT   TRANSMEM   1397   1417       POTENTIAL.
     TRANSMEM   1493   1513       POTENTIAL.
     DOMAIN     1514   1529       CYTOPLASMIC (POTENTIAL).
FT   NP_BIND     920    927       ATP (POTENTIAL).
FT   DOMAIN       13     24       POLY SER
FT   DOMAIN       66     69       POLY SER
FT   DOMAIN      794    797       POLY PHE
FT   CARBOHYD    558    558       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    744    744       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   1529 AA;  172255 MW;  [illegible] CRC64;

  Query Match             64.5%;  Score 40;  DB 1;  Length 1529;
  Best Local Similarity   70.0%;  Pred. No. 10;
  Matches    7;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 ILNNQEGWKK 10
              |||| :|| |
Db        902 ILNNVDGWVK 911


RESULT 12
CDR3_CANAL
ID   CDR3_CANAL     STANDARD;      PRT;  1501 AA.
AC   O42690;
DT   15-JUL-1999 (Rel. 38, Created)
DT   15-JUL-1999 (Rel. 38, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   OPAQUE-SPECIFIC ABC TRANSPORTER CDR3.
     CDR3.
     Candida albicans (Yeast).
     Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX   NCBI_TaxID=5476;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=1006;
RX   MEDLINE=98053831; PubMed=9393682;
RA   Balan I., Alarco A.M., Raymond M.;
RT   "The Candida albicans CDR3 gene codes for an opaque-phase ABC
RT   transporter.";
RL   J. Bacteriol. 179:7210-7218(1997).
CC   -!- FUNCTION: NOT KNOWN.
CC   -!- DEVELOPMENTAL STAGE: REGULATED IN A CELL TYPE-SPECIFIC MANNER.
CC       HIGH LEVELS IN WO-1 OPAQUE CELLS AND UNDETECTABLE LEVELS IN WO-1
CC       WHITE CELLS.
CC   -!- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
CC       (ABC TRANSPORTERS). PDR5 SUBFAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as  long as its content is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; U89714; AAC19886.1; -.
DR   InterPro; IPR001617; -.
DR   Pfam; PF00005; ABC_tran; 2.
DR   PROSITE; PS00211; ABC_TRANSPORTER; 1.
KW   ATP-binding; Transmembrane; Glycoprotein; Transport; Repeat.
FT   DOMAIN        1    502       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    503    523       POTENTIAL.
FT   TRANSMEM    540    560       POTENTIAL.
FT   TRANSMEM    589    609       POTENTIAL.
FT   TRANSMEM    614    634       POTENTIAL.
FT   TRANSMEM    653    673       POTENTIAL.
FT   TRANSMEM    755    775       POTENTIAL.
FT   DOMAIN      776   1175       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM   1176   1196       POTENTIAL.
FT   TRANSMEM   1212   1232       POTENTIAL.
FT   TRANSMEM   1261   1281       POTENTIAL.
FT   TRANSMEM   1297   1317       POTENTIAL.
FT   TRANSMEM   1325   1345       POTENTIAL.
FT   TRANSMEM   1451   1471       POTENTIAL.
FT   NP_BIND     876    883       ATP (POTENTIAL).
SQ   SEQUENCE   1501 AA;  170270 MW;  [illegible] CRC64;

  Query Match             62.9%;  Score 39;  DB 1;  Length 1501;
  Best Local Similarity   70.0%;  Pred. No. 18;
  Matches    7;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 ILNNQEGWKK 10
              |||| :|  |
Db        858 ILNNIDGWVK 867


RESULT 13
GUNA_CALSA
ID   GUNA_CALSA     STANDARD;      PRT;  1742 AA.
AC   P22534;
DT   01-AUG-1991 (Rel. 19, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   ENDOGLUCANASE A PRECURSOR (EC 3.2.1.4) (ENDO-1,4-BETA-GLUCANASE A)
DE   (CELLULASE A).
GN   CELA.
OS   Caldocellum saccharolyticum (Caldicellulosiruptor saccharolyticus).
OC   Bacteria; Firmicutes; Bacillus/Clostridium group;
OC   Thermoanaerobacter group; Caldicellulosiruptor.
OX   NCBI_TaxID=44001;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95336703; PubMed=7612247.
RA   Te'O V.S., Saul D.J., Bergquist P.L.;
RT   "celA, another gene coding for a multidomain cellulase from the
RT   extreme thermophile Caldocellum saccharolyticum.";
RL   Appl. Microbiol. Biotechnol. 43:291-296(1995).
RN   [2]
RP   SEQUENCE OF 1516-1742 FROM N.A.
RX   MEDLINE=91247819; PubMed=2039230.
RA   Leuthi E., Jasmat N.B., Grayling R.A., Love D.R., Bergquist P.L.;
RT   "Cloning, sequence analysis, and expression in Escherichia coli of a
RT   gene coding for a beta-mannanase from the extremely thermophilic
RT   bacterium 'Caldocellum saccharolyticum'.";
RL   Appl. Environ. Microbiol. 57:694-700(1991).
CC   -!- FUNCTION: THE N-TERMINAL DOMAIN OF CELA ENCODES FOR AN
CC       ENDOGLUCANASE ACTIVITY [illegible], THE C-TERMINAL
CC       DOMAIN PROBABLY ACT SYNERGISTICALLY TO HYDROLYZE CRYSTALLINE
CC       CELLULOSE.
CC   -!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF 1,4-BETA-D-GLUCOSIDIC
CC       LINKAGES IN CELLULOSE.
CC   -!- PTM: THE LINKER REGION (ALSO TERMED "HINGE") MAY BE A POTENTIAL
CC       SITE FOR [illegible]
CC   -!- SIMILARITY: IN THE N-TERMINAL SECTION; [illegible] FAMILY
CC       F (FAMILY [illegible] OF GLYCOSYL HYDROLASES).
CC   -!- SIMILARITY: IN THE C-TERMINAL SECTION; [illegible]
CC       L (FAMILY 48 OF GLYCOSYL HYDROLASES).
CC   -
CC   This SWISS-PROT entry is copyright. [illegible]
CC   between the Swiss Institute [illegible]
CC   the European Bioinformatics Institute. [illegible]
CC   use by non-profit institutions [illegible]
CC   modified and this statement is not removed. [illegible]
CC   entities requires a license agreement (See http://www. [illegible]
CC   or send an email to [illegible]
CC
DR   EMBL; L32742; AAA91[illegible]
DR   EMBL; M36063; AAA7[illegible]
DR   EMBL; [illegible]; NOT_ANNOTATED_CDS; [illegible]
DR   PIR; A43745; A43745.
DR   HSSP; P26221; [illegible]
DR   InterPro; IPR00[illegible]
DR   InterPro; IPR0[illegible]
DR   InterPro; IPR[illegible]
DR   Pfam; PF00942; [illegible]
DR   Pfam; PF02011; Glyco_hydro_48; [illegible]
DR   Pfam; PF00759; Glyco_hydro_9; [illegible]
DR   PRINTS; PR00844; GLHYDRLASE48.
DR   PROSITE; PS00[illegible]; GH[illegible]
DR   PROSITE; PS00[illegible]; GH[illegible]
KW   Cellulose degradation; Hydrolase; [illegible]
FT   SIGNAL        1    [illegible]       POTENTIAL.
FT   CHAIN        [illegible]  1742       ENDOGLUCANASE A.
FT   DOMAIN       [illegible]             CATALYTIC.
FT   DOMAIN       [illegible]             LINKER ("HINGE") [illegible]
FT   DOMAIN       [illegible]             CELLULOSE-BINDING [illegible]
FT   DOMAIN       [illegible]             LINKER ("HINGE") [illegible]
FT   DOMAIN       [illegible]             CELLULOSE-BINDING [illegible] SIMILARITY).
FT   DOMAIN       [illegible]             LINKER ("HINGE") [illegible]
FT   DOMAIN       1118  1742              CATALYTIC.
FT   ACT_SITE     [illegible]             BY SIMILARITY.
FT   ACT_SITE     [illegible]             BY SIMILARITY.
FT   ACT_SITE     [illegible]             BY SIMILARITY.
FT   CONFLICT     1545  154[illegible]
SQ   SEQUENCE   1742 AA;  [illegible] MW;  [illegible]

  Query Match   [illegible]
  Best Local Similarity   [illegible]
  Matches    6;  Conservative   [illegible]  Mismatches   [illegible]

Qy          1 LENNQEGW 8
                [illegible]
Db        501 IVNNQSGW 50[illegible]


RESULT 14
MYBH_DICDI
ID   MYBH_DICDI     STANDARD;      PRT;   [illegible] AA.
AC   P34127;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   MYB-LIKE PROTEIN [illegible] (FRAGMENT).
GN   MYBA.
OS   Dictyostelium discoideum (Slime mold).
OC   Eukaryota; Mycetozoa; [illegible]
OX   NCBI_TaxID=44689;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92195[illegible]; PubMed=[illegible]
RA   Stober-Grasser [illegible]
RA   Lipsick J.S.;
RL   "The Myb DNA-binding [illegible]
RL   discoideum.";
```

```
RL   [illegible]
CC   -!- FUNCTION: MAY CONTROL [illegible] DIFFERENTIATION.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (PROBABLE).
CC   -!- SIMILARITY: BELONGS TO THE MYB FAMILY.
CC   -!- [illegible] DOMAIN WHICH [illegible] MB1 [illegible] MB4
CC       [illegible] THE [illegible]
CC
CC   This SWISS-PROT entry is copyright. [illegible]
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation [illegible]
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use [illegible]
CC   modified and this statement is not removed. [illegible]
CC   entities requires a license agreement (See http://www. [illegible]
CC   or send an email to license@isb-sib.ch).
CC
DR   EMBL; [illegible]; NOT_ANNOTATED_CDS; [illegible]
DR   HSSP; [illegible]
DR   DictyDb; [illegible]; mybA.
DR   InterPro; IPR[illegible]
DR   Pfam; [illegible]; DNA-binding; [illegible]
DR   PROSITE; [illegible]; MYB_[illegible]
DR   PROSITE; [illegible]; MYB_[illegible]
DR   PROSITE; [illegible]; MYB_[illegible]
KW   Nuclear protein; DNA-binding; Repeat.
FT   DNA_BIND     [illegible]             MYB.
FT   DNA_BIND     [illegible]             MYB.
FT   DNA_BIND     [illegible]             MYB.
FT   DOMAIN       [illegible]             ASN-RICH.
FT   DOMAIN       [illegible]             ASN-RICH.
FT   DOMAIN       [illegible]             [illegible]
FT   DOMAIN       [illegible]             POLY-ASN.
FT   DOMAIN       [illegible]             POLY-[illegible]
FT   DOMAIN       [illegible]             POLY-ASN.
FT   DOMAIN       [illegible]             POLY-ASN.
FT   DOMAIN       [illegible]             POLY-THR.
FT   DOMAIN       [illegible]             POLY-ASN.
FT   NON_TER      [illegible]
SQ   SEQUENCE   [illegible] AA;  [illegible] MW;  [illegible] CRC64;

  Query Match   [illegible]
  Best Local Similarity   [illegible]  Pred. No.   [illegible]
  Matches   [illegible]  Conservative   [illegible]  Mismatches   [illegible]  Indels   [illegible]  Gaps   [illegible]

Qy        [illegible] NNQEWKR [illegible]

Db        [illegible] NNKNWKR [illegible]


RESULT 15
[illegible]_HUMAN
ID   [illegible]_HUMAN     STANDARD;      PRT;   1[illegible] AA.
AC   [illegible]
DT   [illegible] (Rel. [illegible], Created)
DT   [illegible] (Rel. [illegible], Last sequence update)
DT   [illegible] (Rel. [illegible], Last annotation update)
DE   [illegible] CHAIN ([illegible] TUBULIN).
GN   [illegible]
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=[illegible]; PubMed=[illegible];
RA   [illegible]
RT   "[illegible] human centrosomal [illegible]
RT   [illegible] structure and function.";
RL   [illegible]
CC   -!- [illegible] WITH [illegible]
CC       MATERIA[illegible]
CC   -!- SIMILARITY: [illegible] THE [illegible] FAMILY.
CC
```

```
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; AF201334; AAF09585.1; -.
KW   Microtubules; GTP-binding.
FT   NP_BIND     148    154       GTP (POTENTIAL).
SQ   SEQUENCE   475 AA;  52931 MW;  3E8E717[illegible] CRC64;

  Query Match             61.3%;  Score 34;  DB 1;  Length 175;
  Best Local Similarity   100.0%;  Pred. No. [illegible]
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        4 NQEGWK 9
            ||||||
Db      376 NQEGWK 381

Search completed: August 22, 2001, 14:33:5[illegible]
Job time: 412 sec
```

```
                         GenCore version 4.5
              Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on:          August 22, 2001, 14:27:28  Search time 112.12 Seconds
                                              (without alignments)
                                              12.980 Million cell updates/sec

Title:           US-08-659-254-8
Perfect score:   62
Sequence:        1 ILNNQEGWKKE 11

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        425026 seqs, 132205027 residues

Total number of hits satisfying chosen parameters:       425026

    mum DB seq length: 0
    mum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       SPTREMBL_16:*
                1:  sp_archea:*
                2:  sp_bacteria:*
                3:  sp_fungi:*
                4:  sp_human:*
                5:  sp_invertebrate:*
                6:  sp_mammal:*
                7:  sp_mhc:*
                8:  sp_organelle:*
                9:  sp_phage:*
                10:  sp_plant:*
                11:  sp_rodent:*
                12:  sp_unclassified:*
                13:  sp_vertebrate:*
                14:  sp_virus:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed
     and is derived by analysis of the total score distribution.

                              SUMMARIES
            %
 t.       Query
 o.  Score Match Length DB ID                            Description
---------------------------------------------------------------------------
   1    51  82.3    285  6  Q9GMD0                      q9gmd0 ovis aries
   2    45  72.6    281 13  Q9DG09                      q9dg09 gallus gal
   3    42  67.7    289 13  Q9DG08                      q9dg08 xenopus la
   4    42  67.7    783  2  Q45554                      q45554 bacillus sp
   5    42  67.7    824  2  Q9F216                      q9f216 bacillus sp
   6    41  66.1    800  1  Q59066                      q59066 methanococ
   7    40  64.5    704  3  Q9P811                      q9p811 candida gla
   8    40  64.5   1499  3  Q9UVV5                      q9uvv5 candida gla
   9    40  64.5   1542  3  O74208                      o74208 candida gla
  10    39  62.9    102  4  Q9H2I4                      q9h2i4 homo sapien
  11    39  62.9    396  5  P92121                      p92121 hexamita in
  12    39  62.9    847  4  O16584                      o16584 homo sapien
  13    39  62.9    850 11  Q9JJ15                      q9jj15 mus musculu
  14    39  62.9   1022 11  Q9WVQ0                      q9wvq0 mus musculu
  15    39  62.9   1711  2  P96311                      p96311 anaerocellu
  16    38  61.3     95  5  Q9VWN9                      q9vwn9 drosophila
  17    38  61.3    117 10  Q9SET8                      q9set8 witheringia
  18    38  61.3    119  5  O45637                      o45637 caenorhabdi
  19    38  61.3    226  2  Q9KHV7                      q9khv7 [illegible]
  20    38  61.3    285 13  Q9DG10                      q9dg10 brachydanio
  21    38  61.3    403 11  Q9ERI5                      q9eri5 mus musculu
  22    38  61.3    416  4  Q9Y4E2                      q9y4e2 homo sapien
  23    38  61.3    460  4  Q9Y589                      q9y589 homo sapien
  24    38  61.3    460  4  Q9NWZ3                      q9nwz3 homo sapien
  25    37  59.7     48  9  Q38452                      q38452 bacteriopha
  26    37  59.7    235  5  O02442                      o02442 globodera r
  27    37  59.7    292  5  Q9U1D4                      q9u1d4 caenorhabdi
  28    37  59.7    329 14  Q9EIV8                      q9eiv8 edge hill v
  29    37  59.7    329 14  Q9EIV5                      q9eiv5 edge hill v
  30    37  59.7    341  2  Q9RJ47                      q9rj47 streptomyce
  31    37  59.7    346 10  O24417                      o24417 cucurbita m
  32    37  59.7    369  5  O96622                      o96622 dictyosteli
  33    37  59.7    480  5  Q09973                      q09973 caenorhabdi
  34    37  59.7    674  4  Q9NTW0                      q9ntw0 homo sapien
  35    37  59.7    690  9  Q38513                      q38513 bacteriopha
  36    37  59.7    821  2  Q59241                      q59241 bacillus sp
  37    37  59.7    947 11  Q61489                      q61489 mus musculu
  38    37  59.7    950  4  Q9UI53                      q9ui53 homo sapien
  39    37  59.7    950  4  Q9H0D6                      q9h0d6 homo sapien
  40    37  59.7   1049  5  O00851                      o00851 entamoeba h
  41    37  59.7   1171  5  Q9VTW8                      q9vtw8 drosophila
  42    37  59.7   1245  5  O16568                      o16568 caenorhabdi
  43    37  59.7   1289  2  Q9WZ87                      q9wz87 thermotoga
  44    36  58.1    117 10  Q9SET9                      q9set9 witheringia
  45    36  58.1    171  2  Q9KU46                      q9ku46 vibrio chol

                              ALIGNMENTS

RESULT 1
Q9GMD0
ID   Q9GMD0      PRELIMINARY;      PRT;   285 AA.
AC   Q9GMD0;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN.
GN   STAR.
OS   Ovis aries (Sheep).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Caprinae; Ovis
OX   NCBI_TaxID 9940;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   West L.A., Horvat R.D., Roess D.A., Barisas B.G., Juengel J.L.,
RA   Niswender G.D.;
RT   "Steroidogenic Acute Regulatory Protein and Peripheral-Type
RT   Benzodiazepine Receptor Associate at the Mitochondrial Membrane.";
RL   Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF290202; AAG02464.1; -.
SQ   SEQUENCE   285 AA;  31895 MW;  [illegible] CRC64;

  Query Match             82.3%;  Score 51;  DB 6;  Length 285;
  Best Local Similarity   81.8%;  Pred. No. [illegible];
  Matches    9;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 ILNNQEGWKKE 11
              || :|||||||
Db         89 ILKDQEGWKKE 99


RESULT 2
Q9DG09
ID   Q9DG09      PRELIMINARY;      PRT;   281 AA.
AC   Q9DG09;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN.
```

```
GN   STAR.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; [illegible]
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   PubMed=11064258;
RA   Bauer M.P., Bridham J.T., Langenau D.M., [illegible]
RT   "Conservation of steroidogenic acute regulatory (StAR) protein
RT   structure and expression in vertebrates.";
RL   Mol. Cell. Endocrinol. 168:119-125(2000).
DR   EMBL; AF220436; AAG28591.1; -.
SQ   SEQUENCE   281 AA;  [illegible] MW;  [illegible] CRC64;

  [illegible]ery Match            [illegible];  Score 15;  DB 1;  Length [illegible];
  [illegible]t Local Similarity   [illegible];  Pred. No. [illegible];
  Matches    8;  Conservative    1;  Mismatches  [illegible];  Indels  [illegible];  Gaps  [illegible];

Qy          1 ILNNQEGWKKE 11
              ||  :||||:
Db         89 ILGDQEGWKIE 99


RESULT      3
Q9DG08
ID   Q9DG08      PRELIMINARY;      PRT;   [illegible] AA.
AC   Q9DG08;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN (FRAGMENT).
GN   STAR.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; [illegible]
OC   Amphibia; Batrachia; Anura; Mesobatrachia; [illegible]
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=OVARY;
RX   PubMed=11064258;
RA   Bauer M.P., Bridham J.T., Langenau D.M., [illegible]
[illegible]   "Conservation of steroidogenic acute regulatory (StAR) protein
[illegible]   structure and expression in vertebrates.";
RL   Mol. Cell. Endocrinol. 168:119-125(2000).
DR   EMBL; AF220437; AAG28592.1; -.
FT   NON_TER       1      1
FT   NON_TER     280    280
SQ   SEQUENCE   280 AA;  [illegible] MW;  [illegible] CRC64;

  Query Match             [illegible];  Score 1[illegible];  DB 1;  Length [illegible];
  Best Local Similarity   [illegible];  Pred. No. [illegible];
  Matches    7;  Conservative    2;  Mismatches  [illegible];  Indels  [illegible];  Gaps  [illegible];

Qy          1 ILNNQEGWKKE 11
              ||  :|:|||
Db         93 ILGDQDGWKIE 103


RESULT      4
Q45554
ID   Q45554      PRELIMINARY;      PRT;   [illegible] AA.
AC   Q45554;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   ENDOGLUCANASE PRECURSOR (EC [illegible]) [illegible]
DE   (ALKALINE CELLULASE).
OS   Bacillus sp.
```

```
OC   [illegible]
OC   Bacillus/Staphylococcus group; Bacillus.
OX   NCBI_TaxID=[illegible];
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=[illegible];
RA   Miyatake M., Umeda K.;
RL   Submitted (MAY-[illegible]) to the EMBL/GenBank/DDBJ databases.
CC   -!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF 1,4-BETA-D-GLUCOSIDIC
CC       LINKAGES IN CELLULOSE (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO CELLULASE FAMILY A (FAMILY [illegible] OF GLYCOSYL
CC       HYDROLASES).
DR   EMBL; [illegible]; BAA[illegible]; -.
DR   HSSP; [illegible]; 1A3H.
DR   InterPro; IPR[illegible];
DR   Pfam; PF00150; cellulase; 1.
DR   PROSITE; PS00659; GLYCOSYL_HYDROL_F5; UNKNOWN_1.
KW   [illegible]
FT   SIGNAL        1     30       POTENTIAL.
FT   CHAIN        31    [illegible]       ALKALINE CELLULASE.
FT   ACT_SITE    [illegible]    [illegible]       PROTON DONOR (BY SIMILARITY).
FT   ACT_SITE    [illegible]    [illegible]       NUCLEOPHILE (BY SIMILARITY).
SQ   SEQUENCE   [illegible] AA;  [illegible] MW;  [illegible] CRC64;

  Query Match             [illegible];  Score [illegible];  DB [illegible];  Length 7[illegible];
  Best Local Similarity   [illegible];  Pred. No. [illegible];
  Matches    [illegible];  Conservative    [illegible];  Mismatches  [illegible];  Indels  [illegible];  Gaps  [illegible];

Qy          [illegible] LNNQEGWK [illegible]

Db          [illegible] LNNQEGWK [illegible]


RESULT      5
[illegible]
ID   [illegible]      PRELIMINARY;      PRT;   [illegible] AA.
AC   [illegible];
DT   [illegible]-MAR-[illegible] (TrEMBLrel. [illegible], Created)
DT   [illegible]-MAR-[illegible] (TrEMBLrel. [illegible], Last sequence update)
DT   [illegible]-MAR-[illegible] (TrEMBLrel. [illegible], Last annotation update)
DE   CELLULASE.
OS   Bacillus sp.
OC   Bacteria; [illegible]
OC   Bacillus/Staphylococcus group; Bacillus.
OX   NCBI_TaxID=[illegible];
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=KSM-[illegible];
RA   Hakamada Y., Hatada Y., Koike K., Yoshimatsu T., Kawai S.,
RA   Kobayashi T., [illegible];
RT   "Deduced amino acid sequence and possible catalytic residues of a
RT   thermostable, alkaline cellulase from an alkaliphilic Bacillus
RT   strain.";
RL   Biosci. Biotechnol. Biochem. 64:2281-2289(2000).
DR   EMBL; AB[illegible]; BAB[illegible]; -.
SQ   SEQUENCE   [illegible] AA;  [illegible] MW;  [illegible]EA940FFIDF[illegible]; CRC64;

  Query Match             [illegible];  Score 4[illegible];  DB [illegible];  Length 824;
  Best Local Similarity   [illegible];  Pred. No. [illegible];
  Matches    [illegible];  Conservative    [illegible];  Mismatches  [illegible];  Indels  [illegible];  Gaps  [illegible];

Qy          [illegible] ILNNQEGWK [illegible]

Db          [illegible] LNNEEGWK [illegible]


RESULT      6
[illegible]
ID   [illegible]      PRELIMINARY;      PRT;   [illegible] AA.
AC   [illegible];
DT   [illegible]-JUN-1998 (TrEMBLrel. [illegible], Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   HYPOTHETICAL PROTEIN MJ1672.
GN   MJ1672.
OS   Methanococcus jannaschii.
OC   Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;
OC   Methanococcus.
OX   NCBI_TaxID=2190;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=DSM 2661;
RX   MEDLINE=96337999; PubMed 8688087;
RA   Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA   Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA   Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA   Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA   Scott J.L., Geoghagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA   Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA   Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA   Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT   "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT   jannaschii.";
[illegible]   Science 273:1058-1073(1996).
CC   -!- SIMILARITY: TO M.THERMOAUTOTROPHICUM MTH1092 AND M.TUBERCULOSIS
CC       MTCY16B7.19.
DR   EMBL; U67607; AAB99690.1; -.
DR   TIGR; MJ1672; -.
DR   InterPro; IPR002819; -.
DR   Pfam; PF01966; HD; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   800 AA;  93633 MW;  E4CD1[illegible]9C0FB0C CRC64;

  Query Match             66.1%;  Score 4 ;  DB 3;  Length 800;
  Best Local Similarity   54.5%;  Pred. N[illegible];
  Matches    6;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy        1 ILNNQEGWKKE 11
            : ||  || ||
Db      255 LFNNDNGWNKE 265


RESULT    7
Q9P8I1
ID   Q9P8I1        PRELIMINARY;      PRT;   704 AA.
AC   Q9P8I1;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   ABC-TRANSPORTER CDR2 (FRAGMENT).
[illegible]   Candida glabrata (Yeast) (Torulopsis glabrata).
[illegible]   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX   NCBI_TaxID=5478;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=DSY1041;
RA   Sanglard D., Ischer F., Bille J.;
RT   "The role of ABC-transporter genes in high frequency acquisition of
RT   resistance to azole antifungals in Candida glabrata.";
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF251023; AAF64315.1; -.
DR   InterPro; IPR001687; -.
FT   NON_TER       1      1
FT   NON_TER     704    704
SQ   SEQUENCE   704 AA;  80288 MW;  182[illegible]AC72D999 CRC64;

  Query Match             64.5%;  Score 10;  DB 3;  Length 704;
  Best Local Similarity   70.0%;  Pred. No. 79;
  Matches    7;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy        1 ILNNQEGWKK 10
            |||| :|| |
Db      577 ILNNVDGWVK 586


RESULT    8
Q9UVV5
ID   Q9UVV5        PRELIMINARY;      PRT;  1499 AA.
AC   Q9UVV5;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   ATP BINDING CASSETTE TRANSPORTER.
GN   CGCDR1.
OS   Candida glabrata (Yeast) (Torulopsis glabrata).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX   NCBI_TaxID=5478;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC90030;
RX   MEDLINE 20011185; PubMed 10543759;
RA   Sanglard D., Ischer F., Calabrese D., Majcherczyk P.A., Bille J.;
RT   "The ATP Binding Cassette transporter gene CgCDR1 from Candida
RT   glabrata is involved in the resistance of clinical isolates to azole
RT   antifungal agents.";
RL   Antimicrob. Agents Chemother. 43:[illegible]-27[illegible](1999)
DR   EMBL; AF109723; AAF05069.1; -.
DR   InterPro; IPR001687; -.
DR   InterPro; IPR003439; -.
DR   InterPro; IPR003593; -.
DR   Pfam; PF00005; ABC_tran; 2.
DR   PROSITE; PS00211; ABC_TRANSPORTER; UNKNOWN_1.
DR   SMART; SM00382; AAA; 1.
KW   ATP-binding.
SQ   SEQUENCE   1499 AA;  169274 MW;  [illegible]3A4[illegible]D3246 CRC64;

  Query Match             64.5%;  Score 10;  DB 3;  Length 1499;
  Best Local Similarity   70.0%;  Pred. No. 1.7e+02;
  Matches    7;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy        1 ILNNQEGWKK 10
            |||| :|| |
Db      875 ILNNVDGWVK 884


RESULT    9
O74208
ID   O74208        PRELIMINARY;      PRT;  1542 AA.
AC   O74208;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   ATP-BINDING CASSETTE TRANSPORTER
GN   CGR1.
OS   Candida glabrata (Yeast) (Torulopsis glabrata).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; mitosporic Saccharomycetales; Candida
OX   NCBI_TaxID=5478;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN NCCLS84;
RX   MEDLINE=98325436; PubMed 9661006;
RA   Miyazaki H., Miyazaki Y., Geber A., [illegible] Hitchcock C.,
RA   Falconer D.J., Ward D.J., Marsden K., Bennett J.E.;
RT   "Fluconazole resistance associated with drug efflux and increased
RT   transcription of a drug transporter gene, PDH1, in Candida glabrata.";
RL   Antimicrob. Agents Chemother. 4[illegible]:1695-17[illegible](1998).
DR   EMBL; AF046120; AAC31800.1; -.
DR   InterPro; IPR001687; -.
DR   InterPro; IPR003439; -.
DR   InterPro; IPR003593; -.
DR   Pfam; PF00005; ABC_tran; 2.
DR   PROSITE; PS00[illegible]; [illegible]_RANGE_[illegible] [illegible]NEN WN
DR   SMART; SM00[illegible]; AAA; [illegible]
KW   ATP binding
SQ   SEQUENCE  [illegible] AA;  [illegible]5024 MW;  [illegible]AA[illegible] [illegible]


  Query Match             [illegible]%;  Score [illegible];  DB [illegible];  Length [illegible];
  Best Local Similarity   [illegible]%;  Pred. No. [illegible];
  Matches    7;  Conservative [illegible];  Mismatches [illegible];  Indels [illegible];  Gaps [illegible];

Qy          1 ILNNQEGWKK 10
              ||||  |
Db        903 ILNNVEGWVK 91[illegible]


RESULT 10
Q9H214
ID   Q9H214      PRELIMINARY;      PRT;   [illegible] AA.
AC   Q9H214;
DT   01 MAR 2001 (TrEMBLrel. 16, Created)
DT   01 MAR 2001 (TrEMBLrel. 16, Last sequence update)
DT   01 MAR 2001 (TrEMBLrel. 16, Last annotation update)
DE   DC42.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; [illegible]
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Xu X., Yang Y., [illegible]
RT   Submitted (MAY-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF2679[illegible]; AAG[illegible]
SQ   SEQUENCE  102 AA;  [illegible] MW;  A[illegible]


  Query Match             [illegible]%;  Score [illegible];  DB 4;  Length 1[illegible];
  Best Local Similarity   [illegible]%;  Pred. No. [illegible];
  Matches    6;  Conservative [illegible];  Mismatches [illegible];  Indels [illegible];  Gaps [illegible];

Qy          1 ILNNQEGWKK 10
               | |||:||
Db         83 LLLNQQGWKE 9[illegible]


RESULT 11
P92121
ID   P92121      PRELIMINARY;      PRT;   [illegible] AA.
AC   P92121;
DT   01 MAY 1997 (TrEMBLrel. 03, Created)
DT   01 MAY 1997 (TrEMBLrel. 03, Last sequence update)
DT   01 AUG 1998 (TrEMBLrel. 07, Last annotation update)
DE   DYNEIN BETA CHAIN (FRAGMENT).
OS   Hexamita inflata.
OC   Eukaryota; Diplomonadida; Hexamitidae; Hexamitinae; [illegible]
OX   NCBI_TaxID=280[illegible];
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AZ 4;
RA   Keeling P.J., Doolittle W.F.;
RT   Submitted (DEC-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U82547; AAB41014.1; [illegible]
FT   NON_TER       1      1
FT   NON_TER     306    306
SQ   SEQUENCE  306 AA;  [illegible] MW;  [illegible]


  Query Match             [illegible]%;  Score [illegible];  DB [illegible];  Length [illegible];
  Best Local Similarity   [illegible]%;  Pred. No. [illegible];
  Matches    [illegible];  Conservative [illegible];  Mismatches [illegible];  Indels [illegible];  Gaps [illegible];

Qy          4 NQEGWKKE 11
              ||  | |:
Db        259 NQQGWMKE 26[illegible]
```

```
RESULT 12
[illegible]
ID   [illegible]      PRELIMINARY;      PRT;   84[illegible] AA.
AC   [illegible];
DT   01 NOV 1996 (TrEMBLrel. [illegible], Created)
DT   01 NOV 1996 (TrEMBLrel. [illegible], Last sequence update)
DT   01 MAR 2001 (TrEMBLrel. 16, Last annotation update)
DE   PROTEIN KINASE.
GN   MLK3 OR SPRK.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=THYMUS;
RX   MEDLINE=[illegible]; PubMed=[illegible];
RA   Ing Y.L., Leung I.W., [illegible], Lassam N.J.;
RT   "MLK-3: identification of a widely-expressed protein kinase bearing an
RT   SH3 domain and a leucine zipper-basic region domain."
RL   Oncogene [illegible] (1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=[illegible]; PubMed=[illegible];
RA   Gallo K.A., Mark M.R., Scadden D.T., Wang Z., Gu Q., Godowski P.J.;
RT   "Identification and characterization of SPRK, a novel src-homology 3
RT   domain-containing proline-rich kinase with serine/threonine kinase
RT   activity.";
RL   J. Biol. Chem. [illegible] (1994).
CC   [illegible] FAMILY [illegible] KINASES.
DR   EMBL; [illegible]; AAA[illegible]
DR   EMBL; [illegible]; AAA[illegible]
DR   [illegible]
DR   InterPro; IPR[illegible];
DR   InterPro; IPR[illegible];
DR   InterPro; IPR[illegible];
DR   InterPro; IPR[illegible];
DR   Pfam; PF[illegible]; SH3; 1.
DR   Pfam; PF[illegible]; pkinase; 1.
DR   PRINTS; PR[illegible]; SH3DOMAIN.
DR   PROSITE; PS[illegible]; PROTEIN_KINASE_ATP; 1.
DR   PROSITE; PS[illegible]; PROTEIN_KINASE_DOM; 1.
DR   PROSITE; PS[illegible]; PROTEIN_KINASE_ST; 1.
DR   PROSITE; PS[illegible]; SH3; 1.
DR   SMART; SM[illegible]; [illegible]
KW   ATP-binding; Serine/threonine-protein kinase; Transferase.
SQ   SEQUENCE  84[illegible] AA;  [illegible] MW;  [illegible]


  Query Match             [illegible]%;  Score [illegible];  DB [illegible];  Length 84[illegible];
  Best Local Similarity   [illegible]%;  Pred. No. [illegible];
  Matches    [illegible];  Conservative [illegible];  Mismatches [illegible];  Indels [illegible];  Gaps [illegible];

Qy          [illegible] QEGWKKE [illegible]

Db        [illegible] QEGWKKE [illegible]


RESULT 13
[illegible]
ID   [illegible]      PRELIMINARY;      PRT;   [illegible] AA.
AC   [illegible];
DT   [illegible] (TrEMBLrel. [illegible], Created)
DT   [illegible] (TrEMBLrel. [illegible], Last sequence update)
DT   [illegible] MAR [illegible] (TrEMBLrel. [illegible], Last annotation update)
DE   MIXED LINEAGE KINASE [illegible]
GN   MLK3.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=[illegible];
```

```
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129/OLA;
RX   MEDLINE=20354997; PubMed 10894943;
RA   Saridaki A., Ferraz C., Demaille J., [illegible] C., Roux A. F.;
RT   "Genomic sequencing reveals the [illegible] of the [illegible] and [illegible]
RT   genes and their close vicinity to the [illegible] gene on mouse chromosome
RT   19.";
RL   Cytogenet. Cell Genet. 89:85-88(2000).
CC   -!- SIMILARITY: TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR   EMBL; AF155142; AAF73281.1; -.
DR   InterPro; IPR000719; -.
DR   InterPro; IPR001245; -.
DR   InterPro; IPR001452; -.
DR   InterPro; IPR002290; -.
DR   Pfam; PF00018; SH3; 1.
DR   Pfam; PF00069; pkinase; 1
DR   PRINTS; PR00452; SH3DOMAIN.
DR   PRINTS; PR00109; TYRKINASE.
DR   PROSITE; PS00107; PROTEIN_KINASE_ATP; [illegible].
DR   PROSITE; PS50011; PROTEIN_KINASE_DOM; [illegible]
D    PROSITE; PS00108; PROTEIN_KINASE_ST;
     PROSITE; PS50002; SH3; 1.
     SMART; SM00219; TyrKc; 1.
KW   ATP-binding; Kinase; Serine/threonine protein kinase; Transferase;
KW   Tyrosine protein kinase.
SQ   SEQUENCE   850 AA;  93198 MW;  8F026[illegible] 2[illegible]10E CRC64;

  Query Match              62.9%;  Score [illegible];  DB 11;  Length 85[illegible];
  Best Local Similarity   85.7%;  Pred. No.  4e+02;
  Matches    6;  Conservative    1;  Mismatches   [illegible];  Indels   [illegible];  Gaps    [illegible];

Qy        5 QEGWKKE 11
            |||||:|
Db      397 QEGWKRE 403


RESULT  14
Q9WVQ0
ID   Q9WVQ0      PRELIMINARY;      PRT;  [illegible] AA.
AC   Q9WVQ0;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   SPERM TAIL ASSOCIATED PROTEIN
GN   STAP.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
     NCBI_TaxID=10090;
     [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=TESTIS;
RA   Ohuchi J., Arai T., Kon Y., Watanabe [illegible];
RT   "Sperm tail associated protein.";
RL   Submitted (JUL 1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB029919; BAA82514.1; -.
KW   Sperm.
SQ   SEQUENCE   1022 AA;  119400 MW;  1C41[illegible]4AFCD984A CRC64;

  Query Match              62.9%;  Score [illegible];  DB 11;  Length 1022;
  Best Local Similarity   63.6%;  Pred. No.  [illegible].7e+02;
  Matches    7;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy        1 ILNNQEGWKKE 11
            |:||  | | ||
Db      902 IMNNVEQWAKE 912


RESULT  15
P96311
```

```
ID   P96311      PRELIMINARY;      PRT;  [illegible] AA.
AC   P96311;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   ENDOGLUCANASE A (EC 3.2.1.4) (EN[illegible] [illegible] A) (CELLULASE A)
DE   (FRAGMENT).
GN   CELA.
OS   Anaerocellum thermophilum.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC   Anaerocellum group; Anaerocellum.
OX   NCBI_TaxID 31899;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Z-1320;
RA   Zverlov V.;
RL   Submitted (MAR-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: THE N-TERMINAL DOMAIN OF CELA ENCODES FOR AN
CC       ENDOGLUCANASE ACTIVITY ON [illegible] [illegible] TERMINAL
CC       DOMAIN PROBABLY ACT SYNERGISTICALLY TO HYDROLYZE CRYSTALLINE
CC       CELLULOSE (BY SIMILARITY).
CC   -!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF 1,4-BETA-D-GLUCOSIDIC
CC       LINKAGES IN CELLULOSE.
CC   -!- SIMILARITY: THE N-TERMINAL PART BELONGS TO CELLULASE FAMILY E
CC       (FAMILY [illegible] OF GLYCOSYL HYDROLASES).
CC   -!- SIMILARITY: THE C-TERMINAL PART [illegible] CELLULASE FAMILY [illegible]
CC       (FAMILY 48 OF GLYCOSYL HYDROLASES).
DR   EMBL; Z86105; CAB06786.1; -.
DR   HSSP; P26221; 1IF4.
DR   InterPro; IPR[illegible]; -.
DR   InterPro; IPR001701; -.
DR   InterPro; IPR[illegible]56; -.
DR   Pfam; PF00759; Glyco_hydro_9; 1
DR   Pfam; PF00942; CBD_3; 2.
DR   Pfam; PF02011; Glyco_hydro_48; [illegible].
DR   PRINTS; PR00844; GLHYDRLASE48.
DR   PROSITE; PS00592; GLYCOSYL_HYDROL_[illegible]; [illegible].
DR   PROSITE; PS00698; GLYCOSYL_HYDROL_[illegible]; [illegible].
KW   Cellulose degradation; Hydrolase; [illegible]; Signal; Repeat.
FT   NON_TER       1      1
SQ   SEQUENCE   1711 AA;  189979 MW;  [illegible]87CEB9CD0C2[illegible] CRC64;

  Query Match              62.9%;  Score [illegible];  DB 2;  Length 1711;
  Best Local Similarity   75.0%;  Pred. No.  [illegible]e+02;
  Matches    6;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        1 ILNNQEGW 8
            |:||| |
Db      478 IVNNQSGW 485


Search completed: August 22, 2001, 14:[illegible]:29
Job time: 203 sec
```

```
                      GenCore version 1.5
          Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on:          August 22, 2001, 14:27:27   Search time 112.72 Seconds
                                             (without alignments)
                                             10.620 Million cell updates/sec

Title:           US-08-659-254-7
Perfect score:   53
Sequence:        1 AEHGPTCMV 9

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters:      425026

[illegible]mum DB seq length: 0
[illegible]mum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       SPTREMBL_16:*
                1:  sp_archea:*
                2:  sp_bacteria:*
                3:  sp_fungi:*
                4:  sp_human:*
                5:  sp_invertebrate:*
                6:  sp_mammal:*
                7:  sp_mhc:*
                8:  sp_organelle:*
                9:  sp_phage:*
                10:  sp_plant:*
                11:  sp_rodent:*
                12:  sp_unclassified:*
                13:  sp_vertebrate:*
                14:  sp_virus.*

    Pred. No. is the number of results predicted by chance to have a
    score greater than or equal to the score of the result being printed,
    and is derived by analysis of the total score distribution.

                              SUMMARIES
               %
[illegible]t          Query
[illegible]o.  Score  Match Length DB  ID                      Description
-------------------------------------------------------------------------------
    1     53  100.0    285  6  Q9GMD0               q9gmd0 ovis aries
    2     46   86.8    281 13  Q9DG09               q9dg09 gallus gall
    3     42   79.2    285 13  Q9DC10               q9dc10 brachydanio
    4     42   79.2    287 13  Q9DEB4               q9deb4 oncorhynchu
    5     42   79.2    287 13  Q9DE[illegible]      [illegible]
    6     42   79.2    289 13  Q9DG08               q9dg08 xenopus lae
    7     41   77.4    442  2  Q9HTM8               q9htm8 pseudomonas
    8   39.5   74.5    325  6  O97860               o97860 oryctolagus
    9     37   69.8    342  4  O00510               o00510 homo sapien
   10     35   66.0    208 11  Q9JMG4               q9jmg4 mus musculu
   11     35   66.0    273  2  Q9L1I6               q9l1i6 streptomyce
   12     35   66.0    325  5  O46177               o46177 paragonimus
   13     35   66.0    533 11  P97[illegible]20     p97[illegible]20 mus musculu
   14     35   66.0    557  3  O42902               o42902 schizosacch
   15     35   66.0   1077  4  Q9H3[illegible]4     q9h3[illegible]4 homo sapien
   16     35   66.0   1740  2  Q9KWY8               q9kwy8 streptomyce
   17     34   64.2     96  2  O50176               o50176 pseudomonas
   18     34   64.2    111  4  Q9H416               q9h416 homo sapien
   19     34   64.2    295  3  Q9HEH0               q9heh0 neurospora
   20     34   64.2    330  2  P72434               P72434 salmonella
   21     34   64.2    330  2  Q9EUH3               q9euh3 salmonella
   22     34   64.2    492  2  Q9PE43               q9pe43 xylella fas
   23     34   64.2    499  1  O26214               o26214 methanobact
   24     34   64.2    500  1  O28844               o28844 archaeoglob
   25     34   64.2    549  1  Q9HPE3               q9hpe3 halobacteri
   26     34   64.2    552  5  Q17541               q17541 caenorhabdi
   27     34   64.2    583  5  Q9GZ[illegible]      [illegible] caenorhabdi
   28     34   64.2   1138 14  O55[illegible]45     o55[illegible]45 [illegible] vir
   29     34   64.2   1655 13  Q9PTY1               q9pty1 paralichthy
   30     34   64.2   2543  5  P90649               p90649 paramecium
   31     34   64.2   4370  4  Q9H[illegible]V5     q9h[illegible] homo sapien
   32     33   62.3    136  4  Q9H4[illegible]8     q9h4[illegible]8 homo sapien
   33     33   62.3    146  2  Q53925               q53925 streptomyce
   34     33   62.3    264  4  Q9H5D1               q9h5d1 homo sapien
   35     33   62.3    291  2  Q9I4K0               q9i4k0 pseudomonas
   36     33   62.3    367 11  Q9QXD2               q9qxd2 mus musculu
   37     33   62.3    416  2  Q9RJJ6               q9rjj6 streptomyce
   38     33   62.3    466 10  Q9SYV7               q9syv7 arabidopsis
   39     33   62.3    497  4  Q9H0B8               q9h0b8 homo sapien
   40     33   62.3    511 10  Q9FE[illegible]      q9fe[illegible] oryza sativ
   41     33   62.3    523  3  Q9UTQ7               q9utq7 schizosacch
   42     33   62.3    536  4  Q9U[illegible]12     q9u[illegible]12 homo sapien
   43     33   62.3    611  1  P72112               P72112 nitrosomona
   44     33   62.3    614  5  Q9[illegible]70      q9[illegible]70 drosophila
   45     33   62.3    649  3  Q07655               q07655 saccharomyc

                              ALIGNMENTS

RESULT    1
Q9GMD0
ID   Q9GMD0        PRELIMINARY;      PRT;   285 AA.
AC   Q9GMD0;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN.
GN   STAR.
OS   Ovis aries (Sheep).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Caprinae; Ovis.
OX   NCBI_TaxID 9940;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   West L.A., Horvat R.D., Roess D.A., Barisas B.G., Juengel J.L.,
RA   Niswender G.D.;
RT   "Steroidogenic Acute Regulatory Protein and Peripheral-Type
RT   Benzodiazepine Receptor Associate at the Mitochondrial Membrane.";
RL   Submitted (JUN-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF290202; AAG02464.1; -.
SQ   SEQUENCE   285 AA;  31895 MW;  AB[illegible] CRC64;

  Query Match             100.0%;  Score 53;  DB 6;  Length 285;
  Best Local Similarity   100.0%;  Pred. No. [illegible];
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AEHGPTCMV 9
              |||||||||
Db        218 AEHGPTCMV 226


RESULT    2
Q9DG09
ID   Q9DG09        PRELIMINARY;      PRT;   281 AA.
AC   Q9DG09;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN.
```

```
GN   STAR.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; [illegible]
OC   Archosauria; Aves; Neognathae; [illegible]
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   PubMed 11064158
RA   Bauer M.P., Bradham [illegible] Lanneau [illegible] W.
RT   "Conservation of steroidogenic acute regulatory (StAR) protein
RT   structure and expression in vertebrates."
RL   Mol. Cell. Endocrinol. 168:119-125(2000)
DR   EMBL; AF2204[illegible] AAG2[illegible]
SQ   SEQUENCE   28[illegible] AA;  [illegible] MW;  [illegible]

  Query Match             [illegible]%;  Score 4[illegible]; [illegible]
  Best Local Similarity   [illegible]%;  Pred. No. [illegible]
  Matches    8;  Conservative    1;  Mismatches  [illegible]  Indels  [illegible]  Gaps  [illegible]

Qy          1 AEHGPTCMV 9
              || ||||||
Db        218 AENGPTCMV [illegible]


RESULT     3
Q9DG10
ID   Q9DG10       PRELIMINARY;       PRT;   28[illegible] AA.
AC   Q9DG10;
DT   01 MAR 2001 (TrEMBLrel. 16, Created)
DT   01 MAR 2001 (TrEMBLrel. 16, Last sequence update)
DT   01 MAR 2001 (TrEMBLrel. 16, Last annotation update)
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN
GN   STAR.
OS   Brachydanio rerio (Zebrafish) (Zebra danio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; [illegible]
OC   Actinopterygii; Neopterygii; Teleostei; [illegible]
OC   Cypriniformes; Cyprinidae; Rasborinae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   PubMed 11064158
RA   Bauer M.P., Bradham [illegible] Lanneau [illegible] W.
RT   "Conservation of steroidogenic acute regulatory (StAR) protein
RT   structure and expression in vertebrates.";
RL   Mol. Cell. Endocrinol. 168:119-125(2000)
DR   EMBL; AF22043[illegible] AAG28[illegible]
SQ   SEQUENCE   285 AA;  [illegible] MW;  [illegible]

  Query Match             [illegible]%;  Score 42[illegible]; [illegible]
  Best Local Similarity   [illegible]%;  Pred. No. [illegible]
  Matches    7;  Conservative  [illegible]  Mismatches  [illegible]  Indels  [illegible]  Gaps  [illegible]

Qy          1 AEHGPTCMV 9
              || ||||:|
Db        216 AENGPTCIV 224


RESULT     4
Q9DEB4
ID   Q9DEB4       PRELIMINARY;       PRT;   287 AA.
AC   Q9DEB4;
DT   01 MAR 2001 (TrEMBLrel. 16, Created)
DT   01 MAR 2001 (TrEMBLrel. 16, Last sequence update)
DT   01 MAR 2001 (TrEMBLrel. 16, Last annotation update)
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN
GN   STAR.
OS   Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; [illegible]
OC   Actinopterygii; Neopterygii; Teleostei; [illegible]
OC   Protacanthopterygii; Salmoniformes; Salmonidae; [illegible]
```

```
OX   NCBI_TaxID=[illegible]
RN   [illegible]
RP   SEQUENCE FROM N.A.
RC   TISSUE=HEAD KIDNEY;
RA   [illegible] Kusakabe M., McQuillan [illegible] Young [illegible]
RT   "Molecular cloning of steroidogenic acute regulatory (StAR) protein
RT   cDNA from rainbow trout.";
RL   Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; [illegible]
SQ   SEQUENCE   287 AA;  [illegible] MW;  [illegible] CRC64;

  Query Match             [illegible]%;  Score 4[illegible]; [illegible]  Length 287;
  Best Local Similarity   [illegible]%;  Pred. No. [illegible]
  Matches   [illegible]  Conservative  [illegible]  Mismatches  [illegible]  Indels  [illegible]  Gaps  [illegible]

Qy          1 AEHGPTCMV 9

Db        [illegible] AEN[illegible]V [illegible]


RESULT     [illegible]
[illegible]
ID   [illegible]      PRELIMINARY;       PRT;   28[illegible] AA.
AC   [illegible]
DT   [illegible] MAR [illegible] (TrEMBLrel. [illegible] Created)
DT   01 MAR 200[illegible] (TrEMBLrel. 1[illegible] Last sequence update)
DT   [illegible] MAR [illegible] (TrEMBLrel. [illegible] Last annotation update)
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN
GN   STAR.
OS   Salvelinus fontinalis (Brook trout) (Brook charr).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC   Protacanthopterygii; Salmoniformes; Salmonidae; Salvelinus.
OX   NCBI_TaxID=[illegible]
RN   [illegible]
RP   SEQUENCE FROM N.A.
RC   TISSUE=BRAIN, TESTIS, HEAD KIDNEY, SKIN AND LIVER;
RA   Goetz F.W.;
RT   "[illegible] StAR and MLN64 [illegible]";
RL   Submitted (FEB-[illegible]) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF[illegible]
SQ   SEQUENCE   [illegible] AA;  [illegible] MW;  [illegible] CRC64;

  Query Match             [illegible]%;  Score 4[illegible]; [illegible]  Length 287;
  Best Local Similarity   [illegible]%;  Pred. No. [illegible]
  Matches   [illegible]  Conservative  [illegible]  Mismatches  [illegible]  Indels  [illegible]  Gaps  [illegible]

Qy          1 AEH[illegible]MV 9
              [illegible]
Db        [illegible] AEN[illegible]V [illegible]


RESULT     [illegible]
[illegible]
ID   [illegible]      PRELIMINARY;       PRT;   [illegible] AA.
AC   [illegible]
DT   01 MAR 2[illegible] (TrEMBLrel. [illegible] Created)
DT   [illegible] MAR [illegible] (TrEMBLrel. 16, Last sequence update)
DT   01 MAR [illegible] (TrEMBLrel. [illegible] Last annotation update)
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN (FRAGMENT).
GN   STAR.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=[illegible]
RN   [illegible]
RP   SEQUENCE FROM N.A.
RC   TISSUE=OVARY;
RX   PubMed 11064158
RA   Bauer M.P., Bradham [illegible] Lanneau [illegible] M., [illegible] Goetz F.W.;
```

```
RT   "Conservation of steroidogenic acute regulatory (StAR) protein
RT   structure and expression in vertebrates.";
RL   Mol. Cell. Endocrinol. 168:119-125(2000).
DR   EMBL; AF220437; AAC28595.1; -.
FT   NON_TER       1      1
FT   NON_TER     289    289
SQ   SEQUENCE   289 AA;  32066 MW;  077211A[illegible]86BD8E CRC64;


  Query Match             79.2%;  Score 42;  DB [illegible];  Length 289;
  Best Local Similarity   87.5%;  Pred. No. 9;
  Matches    7;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        2 EHGPTCMV 9
            |:||||||
Db      223 ENGPTCMV 230


RESULT    7
Q9HTM8
ID   Q9HTM8      PRELIMINARY;      PRT;   442 AA.
AC   Q9HTM8;
     01-MAR-2001 (TrEMBLrel. 16, Created)
     01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
D    01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   PROBABLE OXIDOREDUCTASE.
GN   PA5327.
OS   Pseudomonas aeruginosa.
OC   Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC   Pseudomonas.
OX   NCBI_TaxID=287;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PAO1;
RX   MEDLINE=20437337; PubMed=10984043;
RA   Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrener P.,
RA   Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA   Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RA   Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA   Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA   Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT   "Complete genome sequence of Pseudomonas aeruginosa PAO1, an
RT   opportunistic pathogen.";
RL   Nature 406:959-964(2000).
DR   EMBL; AE004945; AAG08712.1; -.
DR   InterPro; IPR001575; -.
DR   Pfam; PF01565; FAD_binding_4; 1.
SQ   SEQUENCE   442 AA;  49435 MW;  0306[illegible]174C4E3 CRC64;


  uery Match              77.4%;  Score 4[illegible];  DB 2;  Length 442;
  st Local Similarity     75.0%;  Pred. No. 1.4;
  tches    6;  Conservative    1;  Mismatches    1;  Indels    [illegible];  Gaps    [illegible];

Qy        1 AEHGPTCM 8
            ||||| |:
Db      318 AEHGPACL 325


RESULT    8
O97860
ID   O97860      PRELIMINARY;      PRT;   325 AA.
AC   O97860;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   TARTRATE-RESISTANT ACID PHOSHATASE.
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX   NCBI_TaxID=9986;
RN   [1]
RP   SEQUENCE FROM N.A.
```

```
RC   STRAIN=JAPAN WHITE; TISSUE=LONG BONES;
RX   MEDLINE=98421349; PubMed=9753427;
RA   Kobori M., Ikeda Y., Nara H., Kato M., Kumegawa M., Nojima H.,
RA   Kawashima H.;
RT   "Large scale isolation of osteoclast-specific genes by an improved
RT   method involving the preparation of a subtracted cDNA library.";
RL   Genes Cells 3:459-475(1998).
CC   -!- CATALYTIC ACTIVITY: AN ORTHOPHOSPHORIC MONOESTER + H(2)O = AN
CC       ALCOHOL + ORTHOPHOSPHATE (BY SIMILARITY).
DR   EMBL; AB009[illegible]; BAA75919.1; -.
DR   InterPro; IPR000934; -.
DR   InterPro; IPR003266; -.
DR   Pfam; PF02227; PA_phosphatase; 1.
KW   Hydrolase; Iron; Zinc.
SQ   SEQUENCE   325 AA;  36624 MW;  9AE2[illegible]B CRC64;


  Query Match             74.5%;  Score [illegible]5;  DB 6;  Length 325;
  Best Local Similarity   80.0%;  Pred. No. [illegible];
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    1;  Gaps    1;

Qy        1 AEHGPT-CMV 9
            |||||| |:|
Db      212 AEHGPTHCLV 221


RESULT    9
O00510
ID   O00510      PRELIMINARY;      PRT;   342 AA.
AC   O00510;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   HYPOTHETICAL 37.4 KDA PROTEIN.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=PLACENTA, AND PBMC;
RX   MEDLINE=96125[illegible]; PubMed=85431[illegible];
RA   Gualandi G., Frezza D., Scotto A., Franchi E., Cattano S.,
RA   Fruscalzo A., Calef E.;
RT   "Integration of an Epstein-Barr virus episome 3' into the gene
RT   encoding immunoglobulin heavy chain alpha 1 in a lymphoblastoid cell
RT   line.";
RL   Gene 166:221-22[illegible](1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=PLACENTA, AND PBMC;
RA   Gualandi G.;
RL   Submitted (NOV-199[illegible]) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=PLACENTA, AND PBMC;
RA   Gualandi G.;
RL   Submitted (APR-1994) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RC   TISSUE=PLACENTA, AND PBMC;
RA   Frezza D.;
RL   Submitted (FEB-1997) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   SEQUENCE FROM N.A.
RC   TISSUE=PLACENTA, AND PBMC;
RA   Frezza D.;
RL   Submitted (JUN-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; X76785; CAA54180.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   342 AA;  37401 MW;  [illegible]10204 CRC64;
```

```
  Query Match            [illegible]
  Best Local Similarity  [illegible]
  Matches    5;  Conservative  [illegible]  Mismatches  [illegible]

Qy        2 EHGPTC [illegible]
              |||
Db      189 QHGPTC 194


RESULT 10
Q9IMG4
ID   Q9IMG4      PRELIMINARY      [illegible]  AA.
AC   Q9IMG4;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   MRNA, COMPLETE [illegible]
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; [illegible] Vertebrata; [illegible]
OC   Mammalia; Eutheria; [illegible] Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=WHOLE BRAIN;
RX   MEDLINE=20145471; PubMed=[illegible];
RA   Inoue S., Sano H., [illegible] M.;
RT   "Growth suppression of Escherichia [illegible]
RT   mammalian genes with transmembrane or ATPase domains.";
RL   Biochem. Biophys. Res. Commun. [illegible]
DR   EMBL; AB030182; BAA9[illegible]
SQ   SEQUENCE  258 AA; [illegible] MW; [illegible]

  Query Match            [illegible]
  Best Local Similarity  [illegible]
  Matches    [illegible]  Conservative  [illegible]  Mismatches  [illegible]

Qy        2 EHGPTCM [illegible]
              |||| |:
Db      109 EHGPGCL 11[illegible]


RESULT 11
Q9LID6
ID   Q9LID6      PRELIMINARY      [illegible]  AA.
AC   Q9LID6;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   PUTATIVE AMINOTRANSFERASE.
GN   SCL11.02[illegible].
OS   Streptomyces coelicolor.
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Streptomycineae; [illegible]
OX   NCBI_TaxID=1902;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A3(2);
RA   Oliver K., Harris [illegible]
RL   Submitted (FEB-2[illegible]) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A3(2);
RA   Cerdeno A.M., Parkhill [illegible] Barrell B.G. [illegible]
RL   Submitted (FEB-20[illegible]) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A3(2);
RX   MEDLINE=97000351; PubMed=8[illegible];
RA   Redenbach M., Kieser H.M., [illegible]
RA   Kinashi H., Hopwood D.A.;
RT   "A set of ordered cosmids and a detailed genetic [illegible]
RT   the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
```

```
RL   M[illegible]
DR   EMBL; AL1[illegible]; CAB[illegible]
DR   InterPro; IPR[illegible]
DR   Pfam; PF[illegible]; aminotran_4; [illegible]
DR   PROSITE; PS[illegible]; AA_TRANSFER_CLASS_4; [illegible]
KW   Transferase; Aminotransferase.
SQ   SEQUENCE  [illegible] AA; [illegible] MW; [illegible]

  Query Match            [illegible]
  Best Local Similarity  [illegible]
  Matches    [illegible]  Mismatches  [illegible]  Indels  [illegible]  Gaps  [illegible]

Qy        [illegible] MV [illegible]

Db        [illegible] VV [illegible]


RESULT [illegible]
[illegible]
ID   [illegible]      PRELIMINARY      [illegible]  AA.
AC   [illegible]
DT   01-JUN-[illegible] (TrEMBLrel. [illegible], Created)
DT   01-JUN-[illegible] (TrEMBLrel. [illegible], Last sequence update)
DT   01-MAR-20[illegible] (TrEMBLrel. 16, Last annotation update)
DE   PREPROCATHEPSIN [illegible]
OS   Paragonimus westermani.
OC   Eukaryota; Metazoa; Platyhelminthes; Platyhelminthes; Neodermata;
OC   Trematoda; Digenea; Plagiorchiida; [illegible] Paragonimidae;
OC   Paragonimus.
OX   NCBI_TaxID=[illegible];
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Yun D.H., [illegible] Yang H.[illegible] Chae S.Y., Kong Y.;
RL   Submitted (FEB-[illegible]) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U[illegible]; AAF[illegible]
DR   [illegible]
DR   InterPro; IPR[illegible]
DR   InterPro; IPR[illegible]
DR   Pfam; PF[illegible]; Peptidase_[illegible]
DR   PRINTS; PR[illegible]; PAPAIN.
DR   PROSITE; PS[illegible] [illegible]ASE_ASN; [illegible]
DR   PROSITE; PS[illegible] THIOL_PROTEASE_CYS; [illegible]
DR   PROSITE; PS[illegible] THIOL_PROTEASE_HIS; [illegible]
KW   Hydrolase; Signal; Thiol protease.
FT   SIGNAL       [illegible]       POTENTIAL.
FT   CHAIN        [illegible]       CATHEPSIN [illegible]
SQ   SEQUENCE  [illegible] AA; [illegible] MW; [illegible]

  Query Match            [illegible]  Length [illegible]
  Best Local Similarity  [illegible]  Pred. No. 4[illegible]
  Matches    [illegible]  Mismatches  [illegible]  Indels  [illegible]  Gaps  [illegible]

Qy        [illegible] AEHG [illegible] M [illegible]
                   |
Db      [illegible] AEH[illegible]M[illegible]


RESULT [illegible]
[illegible]
ID   [illegible]      PRELIMINARY      [illegible]  AA.
AC   [illegible]
DT   01-MAY-[illegible] (TrEMBLrel. [illegible], Created)
DT   01-MAY-[illegible] (TrEMBLrel. [illegible], Last sequence update)
DT   01-AUG-[illegible] (TrEMBLrel. [illegible], Last annotation update)
DE   [illegible] PROTEIN [illegible]
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=[illegible];
RN   [illegible]
RP   SEQUENCE FROM N.A.
```

```
RC   STRAIN-CD1;
RA   Catalano R.D., Vlad M., Kennedy R.C.;
RL   Submitted (DEC-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; Y09878; CAA71005.1; -.
SQ   SEQUENCE   533 AA;  62039 MW;  [illegible] CRC64;


  Query Match             66.0%,  Score 3.3;  DB 11;  Length 533;
  Best Local Similarity   100.0%;  Pred. No. 73;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 HGPTC 7
              |||||
Db        234 HGPTC 238


RESULT  14
O42902
ID   O42902     PRELIMINARY;      PRT;   557 AA.
AC   O42902; P78753;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
     01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
     PROBABLE ASPARAGINE SYNTHETASE [GLUTAMINE-HYDROLYZING] (EC 6.3.5.4)
DE   (GLUTAMINE-DEPENDENT ASPARAGINE SYNTHETASE).
GN   SPBC119.10.
OS   Schizosaccharomyces pombe (Fission yeast).
OC   Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC   Schizosaccharomycetales; Schizosaccharomycetaceae;
OC   Schizosaccharomyces.
OX   NCBI_TaxID=4896;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=972;
RA   Wood V., Rajandream M.A., Barrell B.G., Skelton J., Churcher C.M.;
RL   Submitted (MAR-1998) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN-PR745;
RX   MEDLINE=98162722; PubMed=9501991;
RA   Yoshioka S., Kato K., Nakai K., Okayama H., Nojima H.;
RT   "Identification of open reading frames in Schizosaccharomyces pombe
RT   cDNAs.";
RL   DNA Res. 4:363-369(1997).
CC   -!- CATALYTIC ACTIVITY: ATP + L-ASPARTATE + L-GLUTAMINE = AMP +
CC       PYROPHOSPHATE + L-ASPARAGINE + L-GLUTAMATE.
CC   -!- PATHWAY: ASPARAGINE BIOSYNTHESIS.
CC   -!- SIMILARITY: THE GATASE DOMAIN BELONGS TO TYPE 2 GLUTAMINE
CC       AMIDOTRANSFERASES.
CC   -!- SIMILARITY: TO OTHER ASN SYNTHETASES.
     -!- CAUTION: REF.2 DIFFERS FROM THAT SHOWN FROM POSITION 312 TO 444
         DUE TO A FRAMESHIFT.
     EMBL; AL022117; CAA17925.1; -.
DR   EMBL; D89101; BAA13764.1; ALT_FRAME.
DR   HSSP; P22106; 1CT9.
DR   InterPro; IPR000583; -.
DR   InterPro; IPR001962; -.
DR   Pfam; PF00310; GATase_2; 1.
DR   Pfam; PF00733; Asn_synthase; 1.
DR   PROSITE; PS00443; GATASE_TYPE_II; 1.
KW   Hypothetical protein; Ligase; Asparagine biosynthesis;
KW   Glutamine amidotransferase.
FT   ACT_SITE      2      2       GATASE (BY SIMILARITY).
SQ   SEQUENCE   557 AA;  63240 MW;  E561E[illegible]A92FBAC CRC64;


  Query Match             66.0%;  Score 3.3;  DB 3;  Length 557;
  Best Local Similarity   83.3%;  Pred. No. 76;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 EHGPTC 7
              ||:|||
Db        109 EHGPAC 114
```

```
RESULT  15
Q9H324
ID   Q9H324     PRELIMINARY;      PRT;  1077 AA.
AC   Q9H324;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   ZINC METALLOENDOPEPTIDASE (FRAGMENT).
GN   ADAMTS10.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Apte S.S.;
RT   "ADAM-TS10: A novel member of the ADAM-TS family containing multiple
RT   thrombospondin type 1 repeats.";
RL   Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF163762; AAG35563.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   1077 AA;  118672 MW;  [illegible] CRC64;


  Query Match             66.0%;  Score 3.3;  DB 4;  Length 1077;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 HGPTC 7
              |||||
Db        914 HGPTC 918



Search completed: August 22, 2001, 14:27:28
Job time: 202 sec
```

```
                         GenCore version 4.5
              Copyright (c) 1993 - 2001 Compugen Ltd.

OM protein - protein search, using sw model

Run on:         August 22, 2001, 14:29:11 ; Search time 101.[illegible] Seconds
                                            (without alignments)
                                            6.540 Million cell updates/sec

Title:          US-08-659-254-8
Perfect score:  62
Sequence:       1 IINNQEGWKKE 11

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters:       412676

[illegible]mum DB seq length: 0
[illegible]mum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :      A_Geneseq_0601:*
               1:  /SIDS8/gcgdata/geneseq/geneseqp/AA[illegible]:*
               2:  /SIDS8/gcgdata/geneseq/geneseqp/AA[illegible]:*
               3:  /SIDS8/gcgdata/geneseq/geneseqp/AA198[illegible].DAT:*
               4:  /SIDS8/gcgdata/geneseq/geneseqp/AA1[illegible].DAT:*
               5:  /SIDS8/gcgdata/geneseq/geneseqp/AA1984.DAT:*
               6:  /SIDS8/gcgdata/geneseq/geneseqp/AA1985.DAT:*
               7:  /SIDS8/gcgdata/geneseq/geneseqp/AA198[illegible].DAT:*
               8:  /SIDS8/gcgdata/geneseq/geneseqp/AA1987.DAT:*
               9:  /SIDS8/gcgdata/geneseq/geneseqp/AA1988.DAT:*
              10:  /SIDS8/gcgdata/geneseq/geneseqp/AA[illegible].DAT:*
              11:  /SIDS8/gcgdata/geneseq/geneseqp/AA1990.DAT:*
              12:  /SIDS8/gcgdata/geneseq/geneseqp/AA199[illegible].DAT:*
              13:  /SIDS8/gcgdata/geneseq/geneseqp/AA1992.DAT:*
              14:  /SIDS8/gcgdata/geneseq/geneseqp/AA1993.DAT:*
              15:  /SIDS8/gcgdata/geneseq/geneseqp/AA1994.DAT:*
              16:  /SIDS8/gcgdata/geneseq/geneseqp/AA1995.DAT:*
              17:  /SIDS8/gcgdata/geneseq/geneseqp/AA1996.DAT:*
              18:  /SIDS8/gcgdata/geneseq/geneseqp/AA1997.DAT:*
              19:  /SIDS8/gcgdata/geneseq/geneseqp/AA1998.DAT:*
              20:  /SIDS8/gcgdata/geneseq/geneseqp/AA1999.DAT:*
              21:  /SIDS8/gcgdata/geneseq/geneseqp/AA2000.DAT:*
              22:  /SIDS8/gcgdata/geneseq/geneseqp/AA[illegible]:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                   SUMMARIES

                   %
Result             Query
  No.    Score     Match Length DB   ID                         Description
-------------------------------------------------------------------------------
     1      62    100.0     11  22  AAY72837                    Mouse steroidogeni
     2      62    100.0    284  20  AAW741[illegible]           Mouse steroidogeni
     3      62    100.0    284  22  AAY728[illegible]           Mouse steroidogeni
     4      57     91.9    284  17  AAW132[illegible]           Human steroidogene
     5      57     91.9    285  17  AAW132[illegible]           Human steroidogene
     6      57     91.9    285  17  AAW132[illegible]           Human steroidogene
     7      57     91.9    285  17  AAW00814                    Human steroidogene
     8      57     91.9    285  17  AAW132[illegible]           Human steroidogene
     9      57     91.9    285  17  AAW132[illegible]           Human steroidogene
    10      57     91.9    285  17  AAW132[illegible]           Human steroidogene
    11      57     91.9    285  22  AAY728[illegible]           Human steroidogeni
    12      52     83.9     12  20  AAW74[illegible]            Mouse steroidogeni
    13      52     83.9     12  22  AAY728[illegible]           Mouse tryptic pept
    14      42     67.7    288  19  AAW60[illegible]            Streptococcus pneu
    15      42     67.7    824  21  AAB23[illegible]            Bacillus sp. KSM S
    16      40     64.5   1511  21  AAB28[illegible]            Yeast transporter
    17      39     62.9    102  20  AAY28314                    pk855_1 secreted p
    18      38     61.3      6  22  AAY72[illegible]            Steroidogenic acut
    19      38     61.3    206  21  AAY92[illegible]            Human cancer assoc
    20      38     61.3    226  20  AAW93[illegible]            A. gelatinovorum A
    21      38     61.3    315  20  AAW30[illegible]            Human RIP-interact
    22      37     59.7    202  21  AAG062[illegible]           Arabidopsis thalia
    23      37     59.7    202  21  AAG27[illegible]            Arabidopsis thalia
    24      37     59.7    216  21  AAG27[illegible]            Arabidopsis thalia
    25      37     59.7    217  21  AAG062[illegible]           Arabidopsis thalia
    26      37     59.7    300  22  AAB61[illegible]            Soybean BES1 homol
    27      37     59.7    346  21  AAY8410[illegible]          Amino acid sequenc
    28      37     59.7    800  8   AAP7042[illegible]         Sequence encoded b
    29      37     59.7    822  13  AAR26[illegible]            Alkaline cellulase
    30      37     59.7    950  21  AAB427[illegible]           Human GFPX OR-2506
    31      36     58.1     77  21  AAY6[illegible]             Protein encoded by
    32      36     58.1    144  18  AAW0[illegible]             Immunogenic type F
    33      36     58.1    167  22  AAB6[illegible]             Moraxella catarrha
    34      36     58.1    189  16  AAR6[illegible]             Flower style speci
    35      36     58.1    189  18  AAW14[illegible]            Tomato S-ribonucle
    36      36     58.1    213  18  AAW1[illegible]             Tomato S-ribonucle
    37      36     58.1    330  19  AAW2[illegible]             Human GRBP protein
    38      36     58.1    330  21  AAY9[illegible]             Human Gr140, a sig
    39      36     58.1    330  21  AAY[illegible]              Amino acid sequenc
    40      36     58.1    431  18  AAW0[illegible]             Immunogenic type F
    41      36     58.1    432  21  AAY7[illegible]             Synthetic botulina
    42      36     58.1    [illegible]  22  AA[illegible]       [illegible] skin hea
    43      36     58.1    432  22  AAB01[illegible]            [illegible] skin hea
    44      36     58.1    [illegible]  15  [illegible]         Cellulase [illegible], A
    45      36     58.1    [illegible]  21  AAB1[illegible]     E. coli proliferat

                                   ALIGNMENTS

RESULT    1
AAY72837
ID   AAY72837 standard; peptide; 11 AA.
XX
AC   AAY72837;
XX
DT   31-MAY-2001  (first entry)
XX
DE   Mouse steroidogenic acute regulatory (StAR) peptide #3.
XX
KW   Mouse; steroidogenic acute regulatory protein; StAR;
KW   gene therapy; steroid hormone dependent disorder; LCAH;
KW   lipoid congenital adrenal hyperplasia; sexual maturation;
KW   infertility; androgen-responsive tumour; precocious puberty;
KW   hypogonadotropic hypogonadism; McCune Albright syndrome;
KW   adrenal-hypoplasia congenita; cytostatic.
XX
OS   Mus sp.
XX
PN   US6194555-B1.
XX
PD   27-FEB-2001.
XX
PF   07-JUN-1996;   96US-0659254.
XX
PR   04-NOV-1995;   95US-0538960.
XX
PA   (UYTE ) UNIV TEXAS TECH HEALTH SCI CENT.
XX
PI   Stocco DM, Clark BJ;
XX
DR   WPI; 2001-21793[illegible]/22.
XX
PT   Novel nucleic acid encoding steroidogenic acute regulatory protein
PT   useful for treating steroid hormone dependent [illegible]
XX
PS   Disclosure; [illegible] claim; [illegible] pp; English.
XX
CC   The invention relates to steroidogenic acute regulatory protein
CC   (StAR) and nucleic acid molecules encoding them. [illegible]
CC   are used for regulating steroidogenesis. [illegible]
CC   proteins are involved in regulating cholesterol transport into
CC   the mitochondria of a cell. The [illegible]
CC   invention are useful in gene therapy for the treatment of steroid
CC   hormone dependent disorders, particularly lipoid congenital
CC   adrenal hyperplasia (LCAH), infertility, sexual maturation,
CC   androgen-responsive tumours, precocious puberty, McCune Albright
CC   syndrome, adrenal hypoplasia congenita and hyp[illegible]
CC   hypogonadism. They are also useful for recombinant preparation
CC   of the steroidogenic acute regulatory protein.
CC   The present sequence is steroidogenic acute regulatory (StAR)
CC   peptide. This peptide corresponds to [illegible]
CC   StAR protein. This peptide is useful for [illegible]
[illegible]   for screening and identifying [illegible] of StAR protein
XX
SQ   Sequence   11 AA;

  Query Match   [illegible]%;  Score [illegible];  DB [illegible];  Length [illegible];
  Best Local Similarity   100.0%;  Pred. No. [illegible];
  Matches   11;  Conservative   [illegible];  Mismatches   [illegible];  Indels   [illegible];  Gaps   [illegible];

Qy          1 [illegible]NNQEGWKK[illegible] 11
              [illegible]
Db          1 [illegible]nnqegwkk[illegible] 11


RESULT   2
AAW74153
ID   AAW74153 standard; Protein; 281 AA.
XX
AC   AAW74153;
XX
DT   05 MAY 1999  (first entry)
XX
DE   Mouse steroidogenic acute regulatory protein.
XX
KW   StAR; steroidogenic acute regulatory protein; McCune Albright syndrome;
KW   steroid hormone dependent disorder; adrenal hypoplasia congenita; LCAH;
[illegible]   lipoid congenital adrenal hyperplasia; infertility; sexual maturation;
[illegible]   androgen responsive tumour; hypogonadism; therapy; [illegible]
XX
OS   Mus sp.
XX
PN   US587[illegible]   A.
XX
PD   16 FEB 1999.
XX
PF   04 NOV 1995; [illegible]
XX
PR   04 NOV 1995; [illegible]
XX
PA   (UYTE ) UNIV TEXAS [illegible]
XX
PI   Clark BJ, [illegible] DM;
XX
DR   WPI; [illegible]
DR   N-PSDB; AAX[illegible] AAX[illegible]
XX
PT   New nucleic acid encoding [illegible]
PT   useful for treatment of steroid hormone dependent [illegible]
PT   lipoid congenital adrenal hyperplasia, infertility, and
PT   androgen responsive tumours.
XX
PS   Claim [illegible]; Column [illegible] 38; [illegible]pp; [illegible]
XX
CC   This sequence is the steroidogenic acute regulatory [illegible] (StAR)
```

```
CC   [illegible] the [illegible] the StAR coding sequence and its product [illegible]
CC   [illegible] may be used as probes and primers for PCR hybridisation
CC   studies and in the treatment of steroid hormone dependent disorders such
CC   as lipoid congenital adrenal hyperplasia (LCAH), infertility, sexual
CC   maturation, androgen responsive tumours, precocious puberty,
CC   McCune Albright syndrome, adrenal hypoplasia congenita and hypogonadism
CC   [illegible] StAR may also indicate reduced [illegible]
CC   and [illegible]
XX
SQ   Sequence   284 AA;

  Query Match   [illegible]%;  Score [illegible];  DB [illegible];  Length 284;
  Best Local Similarity   [illegible]%;  Pred. No. [illegible];
  Matches   [illegible];  Conservative   [illegible];  Mismatches   [illegible];  Indels   [illegible];  Gaps   [illegible];

Qy          1 [illegible]NN[illegible]GWKK[illegible]
              [illegible]
Db         [illegible] pegwkk[illegible]


RESULT   [illegible]
AAY[illegible]
ID   AAY[illegible] standard; [illegible]; 284 AA.
XX
AC   AAY[illegible];
XX
DT   [illegible] MAY [illegible] (first entry)
XX
DE   Mouse steroidogenic acute regulatory (StAR) protein.
XX
KW   Mouse; steroidogenic acute regulatory protein; StAR;
KW   gene therapy; steroid hormone dependent disorder; LCAH;
KW   [illegible] adrenal hyperplasia; sexual maturation;
KW   infertility; androgen responsive tumour; precocious puberty;
KW   [illegible]; McCune Albright syndrome;
KW   adrenal hypoplasia congenita; hypogonadism.
XX
OS   Mus sp.
XX
FH   Key                 Location/Qualifiers
FT   Peptide             1..4[illegible]
FT                       /label= Signal peptide
FT   Cleavage-site       [illegible]
FT                       /note= "Site for mitochondrial proteases,
FT                       matrix processing protease and mitochondrial
FT                       intermediate processing peptide"
FT   Protein             [illegible]..284
FT                       /label= Mature steroidogenic acute regulatory protein
FT   Cleavage-site       [illegible]
FT                       /note= "Site for mitochondrial protease,
FT                       matrix processing protease and mitochondrial
FT                       intermediate processing peptide"
FT   Modified-site       [illegible]
FT                       /note= "PKA phosphorylation site"
FT   Modified-site       [illegible]
FT                       /note= "[illegible] phosphorylation site"
FT   Region              [illegible]
FT                       /label= tryptic peptide #[illegible]
FT   Region              [illegible]
FT                       /note= "This sequence is used to raise antibodies
FT                       against StAR protein."
FT   Region              [illegible]
FT                       /label= tryptic peptide #[illegible]
FT   Modified-site       [illegible]
FT                       /note= "PKA phosphorylation site"
FT   Region              [illegible]
FT                       /label= tryptic peptide #[illegible]
FT   Modified-site       [illegible]
FT                       /note= "[illegible] Kinase [illegible] phosphorylation site"
FT   Modified-site       [illegible]
FT                       /note= "PKA phosphorylation site"
XX
```

```
PN   US6194555 B1.
XX
PD   27-FEB-2001.
XX
PF   07-JUN-1996;   96US-0659254.
XX
PR   04-NOV-1995;   95US-0538960.
XX
PA   (UYTE-) UNIV TEXAS TECH HEALTH SCI CEN.
XX
PI   Stocco DM,  Clark BJ;
XX
DR   WPI; 2001-217926/22.
DR   N-PSDB; AAD02901, AAD02902.
XX
PT   Novel nucleic acid encoding steroidogenic acute regulatory protein,
PT   useful for treating steroid hormone-dependent disorders -
XX
PS   Example 2; Fig 2; 41pp; English.
XX
CC   The invention relates to steroidogenic acute regulatory protein
CC   (StAR) and nucleic acid molecules encoding them. StAR proteins
     are used for regulating steroidogenesis. More particularly these
     proteins are involved in regulating cholesterol transport into
CC   the mitochondria of a cell. The nucleic acid molecules of the
CC   invention are useful in gene therapy for the treatment of steroid
CC   hormone-dependent disorders, particularly, lipoid congenital
CC   adrenal hyperplasia (LCAH), infertility, sexual maturation,
CC   androgen-responsive tumours, precocious puberty, McCune-Albright
CC   syndrome, adrenal-hypoplasia congenita or hypogonadotropic
CC   hypogonadism. They are also useful for recombinant preparation
CC   of the steroidogenic acute regulatory protein.
CC   The present sequence is mouse steroidogenic acute regulatory
CC   protein (StAR).
XX
SQ   Sequence   284 AA;


  Query Match             100.0%;  Score 11;  DB 22;  Length 284;
  Best Local Similarity   100.0%;  Pred. No. 0.0056;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ILNNQEGWKKE 11
                ||||||||||| 
Db         88 ilnnqegwkke 98


RESULT    4
AAW13238
ID   AAW13238 standard; Protein; 284 AA.

     AAW13238;

DT   21-MAY-1997  (first entry)
XX
DE   Human steroidogenesis acute regulatory protein mutant delta Arg272.
XX
KW   Human; steroidogenesis; acute regulatory protein; hStAR; analysis;
KW   mutation; detection; prenatal; genetic defect; congenital; protein;
KW   lipoid adrenal hyperplasia; treatment; prevention; gene;
KW   replacement therapy; hypercholesterolaemia; mutant.
XX
OS   Homo sapiens.
XX
PN   WO9629338-A1.
XX
PD   26-SEP-1996.
XX
PF   22-MAR-1996;   96WO-US03896.
XX
PR   23-MAR-1995;   95US-0410540.
XX
PA   (REGC ) UNIV CALIFORNIA.
PA   (UYPE-) UNIV PENNSYLVANIA.
XX
PI   Lin D,  Miller WL,  Strauss JF;
XX
DR   WPI; 1996-443130/44.
XX
PT   Isolated human steroidogenesis acute regulatory protein gene - used
PT   for detection of mutation(s) of this gene that cause congenital
PT   lipoid adrenal hyperplasia
XX
PS   Claim 16; Page -; 89pp; English.
XX
CC   The present sequence is the human steroidogenesis acute regulatory
CC   protein (hStAR), congenital lipoid adrenal hyperplasia (CAH)
CC   associated mutant, delta Arg272 (i.e. wild type residue Arg272 has
CC   been deleted). Analysing the hStAR gene for such a mutation is
CC   useful in the detection (e.g. prenatally) of CAH associated
CC   genetic defects, or its transmission to children. CAH can be
CC   treated by protein or gene replacement therapy, which can also be
CC   used to prevent or treat hypercholesterolaemia.
XX
SQ   Sequence   284 AA;


  Query Match             91.9%;  Score [illegible];  DB 17;  Length 284;
  Best Local Similarity   90.9%;  Pred. No. [illegible];
  Matches   10;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ILNNQEGWKKE 11
                ||:||||||||
Db         89 ilsnqegwkke 99


RESULT    5
AAW13236
ID   AAW13236 standard; Protein; 285 AA.
XX
AC   AAW13236;
XX
DT   21-MAY-1997  (first entry)
XX
DE   Human steroidogenesis acute regulatory protein mutant Ala218Val.
XX
KW   Human; steroidogenesis; acute regulatory protein; hStAR; analysis;
KW   mutation; detection; prenatal; genetic defect; congenital; protein;
KW   lipoid adrenal hyperplasia; treatment; prevention; gene;
KW   replacement therapy; hypercholesterolaemia; mutant.
XX
OS   Homo sapiens.
XX
FH   Key                 Location/Qualifier
FT   Misc-difference 218
FT                       /note= "wild type Ala substituted with Val"
XX
PN   WO9629338-A1.
XX
PD   26-SEP-1996.
XX
PF   22-MAR-1996;   96WO-US03896.
XX
PR   23-MAR-1995;   95US-0410540.
XX
PA   (REGC ) UNIV CALIFORNIA.
PA   (UYPE-) UNIV PENNSYLVANIA.
XX
PI   Lin D,  Miller WL,  Strauss JF;
XX
DR   WPI; 1996-443130/44.
DR   N-PSDB; AAT62021.
XX
PT   Isolated human steroidogenesis acute regulatory protein gene - used
PT   for detection of mutation(s) of this gene that cause congenital
PT   lipoid adrenal hyperplasia
```

```
XX
PS   Claim 13; Page [illegible]pp; English.
XX
CC   The present sequence [illegible]
CC   protein (hStAR) [illegible]
CC   associated mutant, Ala[illegible]Val. Analysing the hStAR gene [illegible]
CC   mutation is useful [illegible] the detection [illegible]
CC   associated genetic defects, or [illegible]
CC   can be treated by protein or gene replacement therapy [illegible]
CC   also be used to prevent or treat hyperch[illegible]
XX
SQ   Sequence   [illegible] AA;

  Query Match [illegible]
  Best Local Similarity [illegible]; Pred. No. [illegible]
  Matches   10; Conservative [illegible]; Mismatches [illegible]

Qy        1 [illegible]NNQF[illegible]WKK[illegible] 11
            [illegible]
Db       89 [illegible]snqeawkk[illegible]


RESULT   6
AAW13237
ID   AAW13237 standard; Protein; [illegible] AA.
XX
AC   AAW13237;
XX
DT   21 MAY 1997  (first entry)
XX
DE   Human steroidogenesis acute regulatory protein [illegible]
XX
KW   Human; steroidogenesis acute regulatory protein; [illegible]
KW   mutation; detection; prenatal; genetic defect; [illegible]
KW   lipoid adrenal hyperplasia; treatment; prevention; [illegible]
KW   replacement therapy; hypercholesterolaemia; [illegible]
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc difference [illegible]
FT                   /note= "wild type [illegible]"
XX
PN   WO9629338-A1.
XX
PD   26 SEP 1996.
XX
PF   22 MAR 1996;  96WO-[illegible]
XX
PR   23 MAR 1995;  [illegible]
XX
PA   (REGC ) UNIV CALIFORNIA.
PA   (UYPE ) UNIV PENNSYLVANIA.
XX
PI   Lin D,  Miller WL, [illegible]
XX
DR   WPI; 1996-44[illegible]
DR   N-PSDB; AAT62022.
XX
PT   Isolated human steroidogenesis acute regulatory protein [illegible]
PT   for detection of mutation(s) of this gene that [illegible]
PT   lipoid adrenal hyperplasia
XX
PS   Claim 13; Page [illegible]; 89pp; English.
XX
CC   The present sequence is the human steroidogenesis acute regulatory
CC   protein (hStAR), congenital lipoid adrenal hyperplasia [illegible]
CC   associated mutant, Leu[illegible]Pro. Analysing the hStAR gene [illegible]
CC   mutation is useful in the detection [illegible] prenatal[illegible]
CC   associated genetic defects, [illegible]
CC   can be treated by protein or gene replacement therapy [illegible]
CC   also be used to prevent or treat hypercholesterolaemia[illegible]
```

```
XX
SQ   Sequence   [illegible] AA;

  Query Match [illegible]
  Best Local Similarity [illegible]; Pred. No. [illegible]
  Matches [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible]

Qy        [illegible]NNQFGWKKF[illegible]

Db        [illegible]qawkk[illegible]


RESULT [illegible]
AAW[illegible]
ID   AAW[illegible] standard; Protein; [illegible] AA.
XX
AC   AAW[illegible];
XX
DT   [illegible] MAY [illegible] (first entry)
XX
DE   Human steroidogenesis acute regulatory protein.
XX
KW   Human; steroidogenesis acute regulatory protein; hStAR; analysis;
KW   mutation; detection; prenatal; genetic defect; congenital; protein;
KW   lipoid adrenal hyperplasia; treatment; prevention; gene
KW   replacement therapy; hypercholesterolaemia.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Region          [illegible]
FT                   /note= "potential site for protein kinase C
FT                   mediated phosphorylation"
FT   Region          [illegible]
FT                   /note= "potential site for protein kinase C
FT                   mediated phosphorylation"
FT   Region          [illegible]
FT                   /note= "potential site for protein kinase A
FT                   mediated phosphorylation"
FT   Region          [illegible]
FT                   /note= "potential site for protein kinase A
FT                   mediated phosphorylation"
FT   Region          [illegible]
FT                   /note= "potential site for protein kinase A
FT                   mediated phosphorylation"
FT   Region          [illegible]
FT                   /note= "potential site for protein kinase C
FT                   mediated phosphorylation"
FT   Region          [illegible]
FT                   /note= "potential site for protein kinase A
FT                   mediated phosphorylation"
FT   Region          [illegible]
FT                   /note= "potential site for protein kinase A
FT                   mediated phosphorylation"
XX
PN   W[illegible]-A[illegible]
XX
PD   [illegible]
XX
PF   [illegible] MAR [illegible]  [illegible]
XX
PR   [illegible] MAR [illegible]  [illegible]
XX
PA   (REGC ) UNIV CALIFORNIA.
PA   (UYPE ) UNIV PENNSYLVANIA.
XX
PI   Lin D,  Miller WL, [illegible]
XX
DR   WPI; 1996-44[illegible]
DR   N-PSDB; AA[illegible]
XX
PT   Isolated human steroidogenesis acute regulatory protein gene - used
PT   for detection of mutation(s) of this gene that cause congenital
PT   lipoid adrenal hyperplasia
XX
PS   Claim 1; Pages 62-63; 89pp; English.
XX
CC   The present sequence is the human steroidogenesis acute
CC   regulatory protein (hStAR). The hStAR gene can be analysed for
CC   mutations to detect (e.g. prenatally) genetic defects associated
CC   with congenital lipoid adrenal hyperplasia (CAH), or its
CC   transmission to children. CAH can be treated by protein or gene
CC   replacement therapy, which can also be used to prevent or treat
CC   hypercholesterolaemia.
CC   A human adrenal cortex cDNA library was screened with a mouse StAR
CC   probe to isolate a 1.6 kb insert, including an ORF for a 285
CC   residue protein. When it was cloned into pSPORT and expressed in
CC   COS-1 cells cotransfected with pP450scc and pADX, it increased the
CC   level of pregnenolone synthesis from cholesterol or
CC   20-alpha-hydroxycholesterol.
XX
SQ   Sequence   285 AA;

   ery Match              91.9%;  Score 57;  DB 17;  Length 285;
   st Local Similarity   90.9%;  Pred. No. [illegible];
  Matches   10;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         1 ILNNQEGWKKE 11
               ||:||||||||
Db        89 ilsnqeqwkke 99


RESULT   8
AAW13233
ID   AAW13233 standard; Protein; 285 AA.
XX
AC   AAW13233;
XX
DT   21-MAY-1997  (first entry)
XX
DE   Human steroidogenesis acute regulatory protein mutant Glu169Gly.
XX
KW   Human; steroidogenesis; acute regulatory protein; hStAR; analysis;
KW   mutation; detection; prenatal; genetic defect; congenital; protein;
KW   lipoid adrenal hyperplasia; treatment; prevention; gene;
KW   replacement therapy; hypercholesterolaemia; mutant.
XX
OS   Homo sapiens.
XX
FH   Key                 Location/Qualifiers
FT   Misc-difference 169
                         /note= "wild type Glu substituted with Gly"

     WO9629338-A1.
XX
PD   26-SEP-1996.
XX
PF   22-MAR-1996;   96WO-US03896.
XX
PR   23-MAR-1995;   95US-0410540.
XX
PA   (REGC ) UNIV CALIFORNIA.
PA   (UYPE-) UNIV PENNSYLVANIA.
XX
PI   Lin D,  Miller WL,  Strauss JF;
XX
DR   WPI; 1996-443130/44.
DR   N-PSDB; AAT62018.
XX
PT   Isolated human steroidogenesis acute regulatory protein gene - used
PT   for detection of mutation(s) of this gene that cause congenital
PT   lipoid adrenal hyperplasia
XX
PS   Claim 13; Page -; 89pp; English.
```

```
XX
CC   The present sequence is the human steroidogenesis acute regulatory
CC   protein (hStAR), congenital lipoid adrenal hyperplasia (CAH)
CC   associated mutant, Glu169Gly. Analysing the hStAR gene for such a
CC   mutation is useful in the detection (e.g. prenatally) of CAH
CC   associated genetic defects, or its transmission to children. CAH
CC   can be treated by protein or gene replacement therapy, which can
CC   also be used to prevent or treat hypercholesterolaemia.
XX
SQ   Sequence   285 AA;

  Query Match             91.9%;  Score 5[illegible];  DB 17;  Length 285;
  Best Local Similarity   90.9%;  Pred. No. [illegible];
  Matches   10;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         1 ILNNQEGWKKE 11
               ||:||||||||
Db        89 ilsnqeqwkke 99


RESULT   9
AAW13234
ID   AAW13234 standard; Protein; 285 AA.
XX
AC   AAW13234;
XX
DT   21-MAY-1997  (first entry)
XX
DE   Human steroidogenesis acute regulatory protein mutant Arg182Leu.
XX
KW   Human; steroidogenesis; acute regulatory protein; hStAR; analysis;
KW   mutation; detection; prenatal; genetic defect; congenital; protein;
KW   lipoid adrenal hyperplasia; treatment; prevention; gene;
KW   replacement therapy; hypercholesterolaemia; mutant.
XX
OS   Homo sapiens.
XX
FH   Key                 Location/Qualifiers
FT   Misc-difference 182
FT                       /note= "wild type Arg substituted with Leu"
XX
PN   WO9629338-A1.
XX
PD   26-SEP-1996.
XX
PF   22-MAR-1996;   96WO-US03896.
XX
PR   23-MAR-1995;   95US-0410540.
XX
PA   (REGC ) UNIV CALIFORNIA.
PA   (UYPE-) UNIV PENNSYLVANIA.
XX
PI   Lin D,  Miller WL,  Strauss JF;
XX
DR   WPI; 1996-443130/44.
DR   N-PSDB; AAT62019.
XX
PT   Isolated human steroidogenesis acute regulatory protein gene - used
PT   for detection of mutation(s) of this gene that cause congenital
PT   lipoid adrenal hyperplasia
XX
PS   Claim 13; Page -; 89pp; English.
XX
CC   The present sequence is the human steroidogenesis acute regulatory
CC   protein (hStAR), congenital lipoid adrenal hyperplasia (CAH)
CC   associated mutant, Arg182Leu. Analysing the hStAR gene for such a
CC   mutation is useful in the detection (e.g. prenatally) of CAH
CC   associated genetic defects, or its transmission to children. CAH
CC   can be treated by protein or gene replacement therapy, which can
CC   also be used to prevent or treat hypercholesterolaemia.
XX
SQ   Sequence   285 AA;
```

```
  Query Match            [illegible]
  Best Local Similarity  [illegible]; Pred. No. [illegible]
  Matches   10; Conservative [illegible]; Mismatches [illegible]

Qy        1 [illegible]NNQ[illegible]WKK[illegible]
              [illegible]
Db       89 ilanqewkk[illegible]


RESULT 10
AAW13235
ID   AAW13235 standard; Protein; 28[illegible] AA.
XX
AC   AAW13235;
XX
DT   21 MAY 1997  (first entry)
XX
DE   Human steroidogenesis [illegible]
XX
KW   Human; steroidogenesis [illegible]
KW   mutation; detection; prenatal; genetic defect; [illegible]
KW   lipoid adrenal hyperplasia; treatment; [illegible]
KW   replacement therapy; hypercholesterolaemia; [illegible]
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference [illegible]
FT                   /note= "wild type [illegible]"
XX
PN   WO96[illegible] A1.
XX
PD   26-SEP-1996.
XX
PF   22-MAR-1996;  96WO-[illegible].
XX
PR   23-MAR-1995;  95US-[illegible].
XX
PA   (REGC ) UNIV CALIFORNIA.
PA   (UYPE ) UNIV PENNSYLVANIA.
XX
PI   Lin D,  Miller WL,  [illegible]
XX
[illegible] WPI; 1996-443[illegible].
[illegible] N-PSDB; AAT6[illegible].
XX
PT   Isolated human steroidogenesis acute regulatory [illegible]
PT   for detection of mutation(s) of [illegible]
PT   lipoid adrenal hyperplasia.
XX
PS   Claim 13; Page [illegible]; [illegible]pp; English.
XX
CC   The present sequence [illegible] the human steroidogenesis acute regulatory
CC   protein (hStAR). Congenital lipoid adrenal hyperplasia (CAH)
CC   associated mutant [illegible]
CC   mutation is useful in the detection [illegible] prenatally, [illegible]
CC   associated genetic defects [illegible]
CC   can be treated by protein or gene replacement therapy [illegible]
CC   also be used to prevent or treat hypercholesterolaemia.
XX
SQ   Sequence   28[illegible] AA;

  Query Match            [illegible]
  Best Local Similarity  [illegible]; Pred. No. [illegible]
  Matches   10; Conservative [illegible]; Mismatches [illegible]; Gaps [illegible]

Qy        1 [illegible]LNNQEGWKKE[illegible]
              [illegible]
Db       89 ilanpeqwkke[illegible]
```

```
RESULT [illegible]
AA[illegible]
ID   AA[illegible] standard; Protein; [illegible] AA.
XX
AC   AA[illegible];
XX
DT   [illegible] MAY [illegible]
XX
DE   Human steroidogenic acute regulatory (StAR) proteins.
XX
KW   Human; steroidogenic acute regulatory protein; StAR;
KW   [illegible] hormone dependent [illegible] StAR;
KW   lipoid congenital adrenal hyperplasia; [illegible] sexual maturation;
KW   [illegible] precocious puberty;
KW   [illegible] McCune Albright syndrome;
KW   adrenal hypoplasia congenital; [illegible]
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference [illegible]
FT                   /note= "[illegible]
FT                   the translation [illegible] the sequence shown
FT                   (AA[illegible])"
FT   Modified-site   [illegible]
FT                   /note= "PKA phosphorylation site"
FT   Modified-site   [illegible]
FT                   /note= "PK phosphorylation site"
FT   Modified-site   [illegible]
FT                   /note= "PKA phosphorylation site"
FT   Modified-site   [illegible]
FT                   /note= "[illegible] kinase [illegible] phosphorylation site"
FT   Modified-site   [illegible]
FT                   /note= "PK phosphorylation site"
FT   Modified-site   [illegible]
FT                   /note= "PKA phosphorylation site"
XX
PN   [illegible]
XX
PD   [illegible]
XX
PF   [illegible]
XX
PR   [illegible] NOV [illegible]
XX
PA   (UYTE ) UNIV TEXAS [illegible]
XX
PI   [illegible] Stark [illegible]
XX
DR   W[illegible]
DR   N-PSDB; AAT[illegible] AAT[illegible].
XX
PT   Novel [illegible] steroidogenic acute regulatory proteins,
PT   useful for treating steroid hormone dependent disorders
XX
PS   Example [illegible]; [illegible]pp; English.
XX
CC   The present [illegible] relates to steroidogenic acute regulatory protein
CC   (StAR) [illegible] encoding them. StAR proteins
CC   are used [illegible] steroidogenesis. More particularly these
CC   proteins are involved in regulating cholesterol transport into
CC   [illegible] the
CC   invention are useful in gene therapy for the treatment of steroid
CC   hormone dependent disorders, particularly lipoid congenital
CC   adrenal hyperplasia (CAH), infertility, sexual maturation
CC   [illegible] precocious puberty, McCune Albright
CC   syndrome, adrenal hypoplasia congenital, hypogonadotropic
CC   hypogonadism. They are also useful for recombinant preparation
CC   of the steroidogenic acute regulatory protein.
CC   The present sequence is human steroidogenic acute regulatory
CC   protein (StAR).
XX
```

```
SQ   Sequence   285 AA;


  Query Match             91.9%;  Score 57;  DB 22;  Length 285;
  Best Local Similarity   90.9%;  Pred. No. 0.039;
  Matches   10;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ILNNQEGWKKE 11
              ||:||||||||
Db         89 ilsnqeqwkke 99


RESULT  12
AAW74155
ID   AAW74155 standard; Protein; 12 AA.
XX
AC   AAW74155;
XX
DT   05-MAY-1999  (first entry)
XX
DE   Mouse steroidogenic acute regulatory protein tryptic peptide.
XX
     StAR; steroidogenic acute regulatory protein; McCune Albright syndrome;
     steroid hormone dependent disorder; adrenal-hypoplasia congenita; LCAH;
KW   lipoid congenital adrenal hyperplasia; infertility; sexual maturation;
KW   androgen-responsive tumours; hypogonadism; therapy; progesterone level.
XX
OS   Mus sp.
XX
PN   US5872230-A.
XX
PD   16-FEB-1999.
XX
PF   04-NOV-1995;   95US-0538960.
XX
PR   04-NOV-1995;   95US-0538960.
XX
PA   (UYTE-) UNIV TEXAS TECH HEALTH SCI CENT.
XX
PI   Clark BJ,  Stocco DM;
XX
DR   WPI; 1999-166720/14
XX
PT   New nucleic acid encoding steroidogenic acute regulatory protein
PT   useful for treatment of steroid hormone dependent disorders such as
PT   lipoid congenital adrenal hyperplasia, infertility and
PT   androgen-responsive tumours
XX
PS   Example 2; Column 26; 25pp; English.
XX
     This sequence is a tryptic peptide fragment of the steroidogenic acute
     regulatory protein (StAR) of the invention. The StAR coding sequence and
CC   its products or derivatives may be used as probes and primers for PCR
CC   hybridisation studies and in the treatment of steroid hormone dependent
CC   disorders such as lipoid congenital adrenal hyperplasia (LCAH),
CC   infertility, sexual maturation, androgen-responsive tumours, precocious
CC   puberty, McCune Albright syndrome, adrenal hypoplasia congenita, and
CC   hypogonadism. Monitoring levels of StAR may also indicate reduced
CC   progesterone levels and so a risk of spontaneous abortion.
XX
SQ   Sequence   12 AA;


  Query Match             83.9%;  Score 52;  DB 20;  Length 12;
  Best Local Similarity   100.0%;  Pred. No. 0.01;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ILNNQEGWK 9
              |||||||||
Db          4 ilnnqeqwk 12


RESULT  13
AAY72833
ID   AAY72833 standard; peptide; 12 AA.
XX
AC   AAY72833;
XX
DT   31-MAY-2001  (first entry)
XX
DE   Mouse tryptic peptide #25 for analysis of 30kDa StAR protein.
XX
KW   Mouse; steroidogenic acute regulatory protein; StAR
KW   gene therapy; steroid hormone-dependent disorder; LCAH;
KW   lipoid congenital adrenal hyperplasia; sexual maturation;
KW   infertility; androgen-responsive tumour; precocious puberty;
KW   hypogonadotropic hypogonadism; McCune Albright syndrome;
KW   adrenal hypoplasia congenita; cytostatic; tryptic peptide.
XX
OS   Mus sp.
XX
PN   US6194555-B1.
XX
PD   27-FEB-2001.
XX
PF   07-JUN-1996;   96US-0659254.
XX
PR   04-NOV-1995;   95US-0538960.
XX
PA   (UYTE-) UNIV TEXAS TECH HEALTH SCI CENT
XX
PI   Stocco DM,  Clark BJ;
XX
DR   WPI; 2001-217946/22.
XX
PT   Novel nucleic acid encoding steroidogenic acute regulatory protein,
PT   useful for treating steroid hormone dependent disorders -
XX
PS   Example 2; Column 28; 41pp; English.
XX
CC   The invention relates to steroidogenic acute regulatory protein
CC   (StAR) and nucleic acid molecules encoding them. StAR proteins
CC   are used for regulating steroidogenesis. More particularly these
CC   proteins are involved in regulating cholesterol transport into
CC   the mitochondria of a cell. The nucleic acid molecules of the
CC   invention are useful in gene therapy for the treatment of steroid
CC   hormone-dependent disorders, particularly, lipoid congenital
CC   adrenal hyperplasia (LCAH), infertility, sexual maturation,
CC   androgen responsive tumours, precocious puberty, McCune Albright
CC   syndrome, adrenal hypoplasia congenita or hypogonadotropic
CC   hypogonadism. They are also useful for recombinant preparation
CC   of the steroidogenic acute regulatory protein.
CC   The present sequence is tryptic peptide #25. This peptide is
CC   used in the cloning and sequence analysis of cDNA encoding the
CC   30kDa form of StAR protein. This peptide is also used in the
CC   analysis of the StAR protein sequence. Peptide #25 is useful
CC   for designing oligonucleotides for screening and identifying
CC   antigenic determinants of StAR protein.
XX
SQ   Sequence   12 AA;


  Query Match             83.9%;  Score 52;  DB 22;  Length 12;
  Best Local Similarity   100.0%;  Pred. No. 0.01;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ILNNQEGWK 9
              |||||||||
Db          4 ilnnqeqwk 12


RESULT  14
AAW60936
ID   AAW60936 standard; Protein; 288 AA.
XX
AC   AAW60936;
```

```
XX
DT   13-OCT-1998  (first entry)
XX
DE   Streptococcus pneumoniae [illegible]
XX
KW   coding region; ORF; [illegible]
KW   infection; prevention; meningitis [illegible]
XX
OS   Streptococcus pneumoniae.
XX
PN   WO981[illegible]-A1
XX
PD   14-MAY-1998.
XX
PF   27-OCT-1997; [illegible] WO-[illegible]
XX
PR   01-NOV-1996; [illegible] GB-[illegible]
XX
PA   (SMIK ) SMITHKLINE BEECHAM [illegible]
PA   (SMIK ) SMITHKLINE BEECHAM [illegible]
XX
PI   Black MT, [illegible]
PI   Reid RH, Zar[illegible]N;
XX
DR   WPI; 1998-28658[illegible]
DR   N-PSDB; AAV[illegible]
XX
PT   New isolated [illegible]
PT   e.g. for identifying anti-bacterial(s) for treatment [illegible]
PT   of meningitis
XX
PS   Claim 13; Page [illegible]; [illegible]pp; English.
XX
CC   The sequence is that of the polypeptide [illegible]
CC   from S. pneumoniae. The protein [illegible]
CC   may be useful in antibacterial [illegible]
CC   prevention of infection, specifically [illegible]
CC   (particularly meningitis) but [illegible]
CC   pylori (ulcers and gastric cancer). It may be [illegible]
CC   use before insertion [illegible]
CC   invasive procedures. The protein [illegible]
CC   it, can also be used [illegible]
CC   and/or humoral immune response [illegible]
CC   antibacterials. The DNA may also [illegible]
CC   that are potential sources of [illegible]
CC   gene expression [illegible]
CC   can be used diagnostically [illegible]
CC   serotyping of [illegible]
XX
SQ   Sequence   268 AA;

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches [illegible]

Qy          1 IINNQEGWKK [illegible]
                [illegible]
Db        260 innnpeqwek [illegible]


RESULT 15
AAB23180
ID   AAB23180 standard; Protein; 824 AA.
XX
AC   AAB23180;
XX
DT   29-JAN-2001  (first entry)
XX
DE   Bacillus sp. KSM [illegible] heat resistant alkaline [illegible]
XX
KW   Heat resistant alkaline cellulase; detergent [illegible]
XX
OS   Bacillus sp. KSM [illegible]
XX
PN   [illegible]-A
XX
PD   [illegible]
XX
PF   [illegible]-JAN-[illegible]
XX
PR   [illegible]-JAN-[illegible]
XX
PA   (KAOS ) KAO [illegible]
XX
DR   WPI; [illegible]
DR   N-PSDB; AAA[illegible]
XX
PT   A heat resistant alkaline cellulase [illegible]
XX
PS   [illegible]
XX
CC   [illegible] heat resistant alkaline cellulase
CC   (AAB23180) from Bacillus sp. KSM [illegible] and the gene encoding it
CC   (AAA[illegible]). [illegible] variants of the heat
CC   resistant alkaline cellulase in which at least one amino acid residue is
CC   deleted, replaced or inserted. The Bacillus sp. KSM [illegible] heat resistant
CC   alkaline cellulase gene was isolated using the Bacillus sp. N[illegible]
CC   alkaline [illegible] derived PCR primers AAA[illegible] and [illegible]. The Bacillus
CC   sp. KSM [illegible] heat resistant alkaline cellulase gene can be used for the
CC   recombinant production of heat resistant alkaline cellulase which can be
CC   formulated in detergent [illegible]. The present sequence
CC   represents the Bacillus sp. KSM [illegible] heat resistant alkaline cellulase
XX
SQ   Sequence   824 AA;

  Query Match [illegible]; Score 4[illegible]; [illegible] Length 824;
  Best Local Similarity [illegible]; Pred. No. 1[illegible];
  Matches [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible]

Qy          [illegible] NNQEWK [illegible]

Db          [illegible] p[illegible]ewk [illegible]



Search completed: August [illegible] 2001, [illegible]
Job time : [illegible]
```

```
                    GenCore vers. [illegible] 1.5
          Copyright (c) 1993 - [illegible] Compugen Ltd.

OM protein - protein search, using sw model

Run on:          August 22, 2001, 14:30:12 ; Search time 5[illegible].29 Seconds
                                            (without alignments)
                                            4.0[illegible]6 Million cell updates/sec

Title:           US-08-659-254-8
Perfect score:   62
Sequence:        1 ILNNQEGWKKE 11

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        197339 seqs, 20590346 residues

Total number of hits satisfying chosen parameters:      197339

[illegible]mum DB seq length: 0
[illegible]mum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       Issued_Patents_AA:*
                1:  /cgn2_6/ptodata/2/[illegible]_COMB.pep:*
                2:  /cgn2_6/ptodata/2/[illegible]_COMB.pep:*
                3:  /cgn2_6/ptodata/2/[illegible]_COMB.pep:*
                4:  /cgn2_6/ptodata/2/[illegible]_COMB.pep:*
                5:  /cgn2_6/ptodata/2/[illegible]_COMB.pep:*
                6:  /cgn2_6/ptodata/2/[illegible].pep:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                SUMMARIES
                   %
Result           Query
  No.    Score   Match Length DB  ID                              Description
 -------------------------------------------------------------------------------
    1       62   100.0     11  2  US-08-538-960-8                 Sequence 8, Appli
    2       62   100.0     11  4  US-08-659-254-8                 Sequence 8, Appli
    3       62   100.0    276  4  US-08-659-254-2                 Sequence 2, Appli
    4       62   100.0    284  2  US-08-538-960-2                 Sequence 2, Appli
    5       57    91.9    285  1  US-08-41[illegible]-2           Sequence 2, Appli
    6       57    91.9    285  4  US-08-659-254-18                Sequence 18, Appl
    7       52    83.9     12  2  US-08-538-960-4                 Sequence 4, Appli
    8       52    83.9     12  4  US-08-659-254-4                 Sequence 4, Appli
    9       38    61.3      6  2  US-08-538-960-6                 Sequence 6, Appli
   10       38    61.3      6  4  US-08-659-254-6                 Sequence 6, Appli
   11       36    58.1    330  2  US-08-81[illegible]-176-1       Sequence 1, Appli
   12       34    54.8     96  2  US-08-48[illegible]-337-12      Sequence 12, Appl
   13       34    54.8     96  2  US-08-37[illegible]-315-12      Sequence 12, Appl
   14       34    54.8     96  5  PCT-US9[illegible]-06552-12     Sequence 12, Appl
   15       34    54.8    228  2  US-08-4[illegible]-395-21       Sequence 21, Appl
   16       34    54.8    333  1  US-08-208-037A-9                Sequence 9, Appli
   17       34    54.8    333  1  US-08-3[illegible]-121B-5       Sequence 5, Appli
   18       34    54.8    333  3  US-08-83[illegible]-526-[illegible]  Sequence [illegible], Appl
   19       34    54.8    333  3  US-09-120-365-65                Sequence 65, Appl
   20       34    54.8    333  4  US-09-515-[illegible]39-65      Sequence 65, Appl
   21       34    54.8    333  5  PCT-US95-1[illegible]820-5      Sequence 5, Appli
   22       34    54.8    334  3  US-08-883-525-1                 Sequence 1, Appli
   23       34    54.8    334  3  US-08-883-525-4                 Sequence 4, Appli
   24       34    54.8    371  1  US-08-74[illegible]-789A-2      Sequence 2, Appli
   25       34    54.8    694  3  US-08-724-[illegible]86-2       Sequence 2, Appli
   26       34    54.8    705  2  US-08-6[illegible]-[illegible]40-2  Sequence 2, Appli
   27       34    54.8    708  2  US-08-6[illegible]5-[illegible]40-4  Sequence 4, Appli
   28       34    54.8    709  1  US-08-[illegible]-2             Sequence 2, Appli
   29       34    54.8    709  1  US-08-4[illegible]-2            Sequence 2, Appli
   30       34    54.8    709  3  US-08-[illegible]-167-2         Sequence 2, Appli
   31       34    54.8    709  3  US-09-[illegible]-2             Sequence 2, Appli
   32       34    54.8    709  4  US-08-[illegible]-798-2         Sequence 2, Appli
   33       34    54.8    711  1  US-08-[illegible]-681-2         Sequence 2, Appli
   34       34    54.8    711  1  US-08-[illegible]-308-2         Sequence 2, Appli
   35       34    54.8    711  1  US-08-[illegible]-029-4         Sequence 4, Appli
   36       34    54.8    711  1  US-08-461-[illegible]-4         Sequence 4, Appli
   37       34    54.8    711  1  US-08-4[illegible]-2            Sequence 2, Appli
   38       34    54.8    711  2  US-08-[illegible]-2             Sequence 2, Appli
   39       34    54.8    711  3  US-08-[illegible]4-167-4        Sequence 4, Appli
   40       34    54.8    711  3  US-09-[illegible]-4             Sequence 4, Appli
   41       34    54.8    711  3  US-08-[illegible]-2             Sequence 2, Appli
   42       34    54.8    711  4  US-08-[illegible]-4             Sequence 4, Appli
   43       34    54.8    711  4  US-09-2[illegible]-577-2        Sequence 2, Appli
   44       34    54.8    711  5  PCT-US9[illegible]-2            Sequence 2, Appli
   45       34    54.8    753  1  US-08-71[illegible]-24[illegible]-2  Sequence 2, Appli

                                ALIGNMENTS

RESULT     1
US-08-538-960-8
; Sequence 8, Application US/0853896[illegible]
; Patent No. 5872[illegible]30
;  GENERAL INFORMATION:
;     APPLICANT:  Stocco, Douglas M.
;     APPLICANT:  Clark, Barbara J.
;     TITLE OF INVENTION:  [illegible] METHODS FOR
;     TITLE OF INVENTION:  [illegible]
;     NUMBER OF SEQUENCES:  14
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Akin, Gump, Strauss, Hauer & Feld, L.L.P.
;       STREET:  [illegible] Bank Plaza, [illegible] Avenue
;       CITY:  Austin
;       STATE:  TX
;       COUNTRY:  U.S.A.
;       ZIP:  78701
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC DOS/MS DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.[illegible]
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  [illegible]
;       FILING DATE:  Concurrently Herewith
;       CLASSIFICATION:  530
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Mayfield, Denise L.
;       REGISTRATION NUMBER:  33,732
;       REFERENCE/DOCKET NUMBER:  4[illegible]/TLM
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  512/499-6200
;       TELEFAX:  512/499-6290
;  INFORMATION FOR SEQ ID NO:  8:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  11 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
US-08-538-960-8


  Query Match             100.0%;  Score 62;  DB 2;  Length 11;
  Best Local Similarity   100.0%;  Pred. No. 2.9e-05;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ILNNQEGWKKE 11
              |||||||||||
Db          1 ILNNQEGWKKE 11
```

```
RESULT 2
US-08-659-254-8
; Sequence 8, Application US/08659254
; Patent No. 6194555
;  GENERAL INFORMATION:
;     APPLICANT: [illegible] Douglas M
;     APPLICANT: Clark, Dr. Barbara
;     TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR REGULATION
;     TITLE OF INVENTION: OF STEROIDOGENESIS
;     NUMBER OF SEQUENCES: [illegible]
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE: Akin, Gump, Strauss, Hauer & Feld, L.L.P.
;       STREET: 1900 Frost Bank Plaza, 816 Congress Avenue
;       CITY: Austin
;       STATE: TX
;       COUNTRY: U.S.A.
;       ZIP: 78701
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE: Floppy disk
;       COMPUTER: IBM PC compatible
;       OPERATING SYSTEM: PC-DOS/MS-DOS
;       SOFTWARE: PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER: US/08/659,254
;       FILING DATE: 07-JUN-1996
;       CLASSIFICATION: [illegible]
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER: US 08/[illegible]
;       FILING DATE: [illegible]-OCT-1995
;     ATTORNEY/AGENT INFORMATION:
;       NAME: Mayfield, Denise L.
;       REGISTRATION NUMBER: [illegible]
;       REFERENCE/DOCKET NUMBER: [illegible]
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE: 512/499-6200
;       TELEFAX: 512/499-[illegible]
;  INFORMATION FOR SEQ ID NO: 8:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH: 11 amino acids
;       TYPE: amino acid
;       STRANDEDNESS: single
;       TOPOLOGY: linear
;     MOLECULE TYPE: peptide
US-08-659-254-8

  Query Match             [illegible]
  Best Local Similarity   [illegible]
  Matches   11;  Conservative  [illegible]  Mismatches  [illegible]  Indels  [illegible]  Gaps  [illegible]

Qy          1 LLNNQEGWKKL 11
              |||||||||||
Db          1 LLNNQEGWKKL 11


RESULT 3
US-08-659-254-2
; Sequence 2, Application US/08659254
; Patent No. 6194555
;  GENERAL INFORMATION
;     APPLICANT: [illegible] Douglas M
;     APPLICANT: Clark, Dr. Barbara
;     TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR REGULATION
;     TITLE OF INVENTION: OF STEROIDOGENESIS
;     NUMBER OF SEQUENCES: [illegible]
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE: Akin, Gump, Strauss, Hauer & Feld, L.L.P.
;       STREET: 1900 Frost Bank Plaza, 816 Congress Avenue
;       CITY: Austin
;       STATE: TX
;       COUNTRY: U.S.A.
;       ZIP: [illegible]
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE: Floppy disk
;       COMPUTER: IBM PC compatible
;       OPERATING SYSTEM: PC-DOS/MS-DOS
;       SOFTWARE: PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER: US/08/659,254
;       FILING DATE: [illegible]-JUN-1996
;       CLASSIFICATION: 530
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER: US 08/[illegible]
;       FILING DATE: [illegible]
;     ATTORNEY/AGENT INFORMATION:
;       NAME: Mayfield, Denise L.
;       REGISTRATION NUMBER: [illegible]
;       REFERENCE/DOCKET NUMBER: [illegible]
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE: [illegible]
;       TELEFAX: [illegible]
;  INFORMATION FOR SEQ ID NO: [illegible]
;     SEQUENCE CHARACTERISTICS:
;       LENGTH: [illegible] amino acids
;       TYPE: amino acid
;       STRANDEDNESS: single
;       TOPOLOGY: linear
;     MOLECULE TYPE: protein
US-08-659-254-2

  Query Match             [illegible]
  Best Local Similarity   [illegible]
  Matches   [illegible]  Conservative  [illegible]  Mismatches  [illegible]  Indels  [illegible]  Gaps  [illegible]

Qy          1 LLNNQEGWKKL 11
              |||||||||||
Db         98 LLNNQEGWKKL 108


RESULT 4
[illegible]
; Sequence [illegible], Application [illegible]
; Patent No. [illegible]
;  GENERAL INFORMATION:
;     APPLICANT: [illegible] Douglas M
;     APPLICANT: Clark, Barbara J
;     TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
;     TITLE OF INVENTION: REGULATION OF STEROIDOGENESIS
;     NUMBER OF SEQUENCES: [illegible]
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE: Akin, Gump, Strauss, Hauer & Feld, L.L.P.
;       STREET: 1900 Frost Bank Plaza, 816 Congress Avenue
;       CITY: Austin
;       STATE: TX
;       COUNTRY: U.S.A.
;       ZIP: [illegible]
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE: Floppy disk
;       COMPUTER: IBM PC compatible
;       OPERATING SYSTEM: PC-DOS/MS-DOS
;       SOFTWARE: PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER: US/08/798,960
;       FILING DATE: [illegible]
;       CLASSIFICATION: [illegible]
;     ATTORNEY/AGENT INFORMATION:
;       NAME: Mayfield, Denise L.
;       REGISTRATION NUMBER: [illegible]
;       REFERENCE/DOCKET NUMBER: [illegible]/ELM
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE: [illegible]
;       TELEFAX: [illegible]
;  INFORMATION FOR SEQ ID NO: [illegible]
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  284 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-538-960-2


  Query Match             100.0%;  Score 5_;  DB 2;  Length 284;
  Best Local Similarity   100.0%;  Pred. No. 0.0011;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ILNNQEGWKKE 11
              |||||||||||
Db         88 ILNNQEGWKKE 98


RESULT    5
US-08-410-540-2
; Sequence 2, Application US/08410540
; Patent No. 5807678
;  GENERAL INFORMATION:
;    APPLICANT:  Miller, Walter L.
;    APPLICANT:  Lin, Dong
;    APPLICANT:  Strauss III, Jerome F
;    TITLE OF INVENTION:  IDENTIFICATION OF GENE MUTATIONS
;    TITLE OF INVENTION:  ASSOCIATED WITH [illegible]
;    NUMBER OF SEQUENCES:  30
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Cooley Godward Castro Huddleson & Tatum
;      STREET:  5 Palo Alto Square
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  US
;      ZIP:  94306-2155
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/410,540
;      FILING DATE:  23-MAR-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Neeley, Richard L.
;      REGISTRATION NUMBER:  30,092
;      REFERENCE/DOCKET NUMBER:  UCAL-238/0[illegible]
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415 853 5070
;      TELEFAX:  415 857 0663
;      TELEX:  380816COOLEYPA
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  285 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-410-540-2


  Query Match             91.9%;  Score 5_;  DB 1;  Length 285;
  Best Local Similarity   90.9%;  Pred. No. 0.0086;
  Matches   10;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ILNNQEGWKKE 11
              ||:||||||||
Db         89 ILSNQEGWKKE 99


RESULT    6
US-08-659-254-18
```

```
; Sequence 18, Application US/08659254
; Patent No. 6194555
;  GENERAL INFORMATION:
;    APPLICANT:  Stocco, Douglas M.
;    APPLICANT:  Clark, Dr. Barbara J.
;    TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR REGULATION
;    TITLE OF INVENTION:  OF STEROIDOGENESIS
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Akin, Gump, Strauss, Hauer & Feld, L.L.P
;      STREET:  1900 Frost Bank Plaza, 816 Congress Avenue
;      CITY:  Austin
;      STATE:  TX
;      COUNTRY:  U.S.A.
;      ZIP:  78701
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.3
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/659,254
;      FILING DATE:  07 JUN 1996
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/538,960
;      FILING DATE:  04-OCT-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Mayfield, Denise L.
;      REGISTRATION NUMBER:  33,732
;      REFERENCE/DOCKET NUMBER:  43375.0006
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  512/499 6200
;      TELEFAX:  512/499-6290
;  INFORMATION FOR SEQ ID NO:  18:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  285 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
US-08-659-254-18


  Query Match             91.9%;  Score 5_;  DB 1;  Length 285
  Best Local Similarity   90.9%;  Pred. No. 0.0086;
  Matches   10;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ILNNQEGWKKE 11
              ||:||||||||
Db         89 ILSNQEGWKKE 99


RESULT    7
US-08-538-960-4
; Sequence 4, Application US/08538960
; Patent No. 5872230
;  GENERAL INFORMATION:
;    APPLICANT:  Stocco, Douglas M.
;    APPLICANT:  Clark, Barbara J.
;    TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR
;    TITLE OF INVENTION:  REGULATION OF STEROIDOGENESIS
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Akin, Gump, Strauss, Hauer & Feld, L.L.P.
;      STREET:  1900 Frost Bank Plaza, 816 Congress Avenue
;      CITY:  Austin
;      STATE:  TX
;      COUNTRY:  U.S.A.
;      ZIP:  78701
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.
;    CURRENT APPLICATION DATA:
;       APPLICATION NUMBER: [illegible]
;       FILING DATE: Concurrently herewith
;       CLASSIFICATION: [illegible]
;    ATTORNEY/AGENT INFORMATION:
;       NAME: Mayfield, Denise L.
;       REGISTRATION NUMBER: [illegible]
;       REFERENCE/DOCKET NUMBER: [illegible] M
;    TELECOMMUNICATION INFORMATION:
;       TELEPHONE: [illegible]
;       TELEFAX: 512/499 [illegible]
;  INFORMATION FOR SEQ ID NO: 4:
;    SEQUENCE CHARACTERISTICS:
;       LENGTH: [illegible] amino acids
;       TYPE: amino acid
;       STRANDEDNESS: single
;       TOPOLOGY: linear
;    MOLECULE TYPE: peptide
 8 538 [illegible] 4

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches [illegible]; Conservative [illegible]; Mismatches [illegible]

Qy          1 LLNNQEGWK [illegible]
              |||||||
Db          4 LLNNQEGWK [illegible]


RESULT 8
US-08-659-254-4
; Sequence 4, Application US/08659254
; Patent No. 6194[illegible]
;  GENERAL INFORMATION:
;       APPLICANT: [illegible] M
;       APPLICANT: Clark, Dr. Barbara
;       TITLE OF INVENTION: COMPOSITIONS AND METHODS [illegible]
;       TITLE OF INVENTION: [illegible] STEROIDOGENESIS
;       NUMBER OF SEQUENCES: 19
;       CORRESPONDENCE ADDRESS:
;          ADDRESSEE: Akin, Gump, Strauss, Hauer & Feld [illegible]
;          STREET: [illegible] Bank Plaza [illegible]
;          CITY: Austin
;          STATE: TX
;          COUNTRY: U.S.A.
;          ZIP: 7870[illegible]
;       COMPUTER READABLE FORM:
;          MEDIUM TYPE: Floppy disk
;          COMPUTER: IBM PC compatible
;          OPERATING SYSTEM: PC-DOS/MS-DOS
;          SOFTWARE: PatentIn Release #1.0, Version #1.[illegible]
;       CURRENT APPLICATION DATA:
;          APPLICATION NUMBER: [illegible]
;          FILING DATE: [illegible]
;          CLASSIFICATION: [illegible]
;       PRIOR APPLICATION DATA:
;          APPLICATION NUMBER: [illegible]
;          FILING DATE: [illegible]
;       ATTORNEY/AGENT INFORMATION:
;          NAME: Mayfield, Denise L.
;          REGISTRATION NUMBER: [illegible]
;          REFERENCE/DOCKET NUMBER: [illegible]
;       TELECOMMUNICATION INFORMATION:
;          TELEPHONE: 512 499 [illegible]
;          TELEFAX: 512/4[illegible]
;  INFORMATION FOR SEQ ID NO: 4:
;       SEQUENCE CHARACTERISTICS:
;          LENGTH: 12 amino acids
;          TYPE: amino acid
;          STRANDEDNESS: single
;          TOPOLOGY: linear
;       MOLECULE TYPE: peptide
```

```
US-[illegible]-4

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible]

Qy          1 LNNQEGWK [illegible]

Db          1 LNNQEGWK [illegible]


RESULT [illegible]
US-[illegible]
; Sequence [illegible], Application US/[illegible]
; Patent No. [illegible]
;  GENERAL INFORMATION:
;       APPLICANT: [illegible] M
;       APPLICANT: Clark, Barbara
;       TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR [illegible]
;       TITLE OF INVENTION: REGULATION OF STEROIDOGENESIS
;       NUMBER OF SEQUENCES: [illegible]
;       CORRESPONDENCE ADDRESS:
;          ADDRESSEE: Akin, Gump, Strauss, Hauer & Feld, L.L.P.
;          STREET: [illegible] Bank Plaza [illegible] Avenue
;          CITY: Austin
;          STATE: TX
;          COUNTRY: U.S.A.
;          ZIP: [illegible]
;       COMPUTER READABLE FORM:
;          MEDIUM TYPE: Floppy disk
;          COMPUTER: IBM PC compatible
;          OPERATING SYSTEM: PC-DOS/MS-DOS
;          SOFTWARE: PatentIn Release #1.0, Version #1.[illegible]
;       CURRENT APPLICATION DATA:
;          APPLICATION NUMBER: [illegible]
;          FILING DATE: Concurrently herewith
;          CLASSIFICATION: [illegible]
;       ATTORNEY/AGENT INFORMATION:
;          NAME: Mayfield, Denise L.
;          REGISTRATION NUMBER: [illegible]
;          REFERENCE/DOCKET NUMBER: [illegible] M
;       TELECOMMUNICATION INFORMATION:
;          TELEPHONE: 512/[illegible]
;          TELEFAX: 512/4[illegible]
;  INFORMATION FOR SEQ ID NO: [illegible]:
;       SEQUENCE CHARACTERISTICS:
;          LENGTH: [illegible] amino acids
;          TYPE: amino acid
;          STRANDEDNESS: single
;          TOPOLOGY: linear
;       MOLECULE TYPE: peptide
US-[illegible]

  Query Match [illegible]
  Best Local Similarity [illegible]; Pred. No. [illegible]
  Matches [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible]

Qy          1 NQEGWK [illegible]
                |
Db          1 NQEGWK [illegible]


RESULT [illegible]
US-[illegible]-4
; Sequence [illegible], Application US/[illegible]
; Patent No. [illegible]
;  GENERAL INFORMATION:
;       APPLICANT: [illegible] M
;       APPLICANT: Clark, [illegible] Barbara
;       TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR REGULATION
;       TITLE OF INVENTION: OF STEROIDOGENESIS
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Akin, Gump, Strauss, Hauer & Feld, L.L.P.
;      STREET:  1900 Frost Bank Plaza, 816 Congress Avenue
;      CITY:  Austin
;      STATE:  TX
;      COUNTRY:  U.S.A.
;      ZIP:  78701
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/659,254
;      FILING DATE:  07-JUN-1996
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/538,960
;      FILING DATE:  04-OCT-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Mayfield, Denise L.
;      REGISTRATION NUMBER:  33,732
;      REFERENCE/DOCKET NUMBER:  43375 0006
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  512/499-6200
;      TELEFAX:  512/499-6290
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  6 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-659-254-6

  Query Match             61.3%;  Score 38;  DB 4;  Length 6;
  Best Local Similarity   100.0%;  Pred. No. 1.5e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       4 NQEGWK 9
           ||||||
Db       1 NQEGWK 6


RESULT  11
US-08-815-176-1
; Sequence 1, Application US/08815176
; Patent No. 5874224
;  GENERAL INFORMATION:
;    APPLICANT:  Bandman, Olga
;    APPLICANT:  DiEgidio, Tony
;    TITLE OF INVENTION:  NOVEL EGF RECEPTOR BINDING PROTEIN
;    NUMBER OF SEQUENCES:  5
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Dr.
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/815,176
;      FILING DATE:  Filed Herewith
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
```

```
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0236 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-855-0555
;      TELEFAX:  415-845-4166
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  330 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    IMMEDIATE SOURCE:
;      LIBRARY:  PGNN0103
;      CLONE:  1291904
US-08-815-176-1

  Query Match             58.1%;  Score 36;  DB 2;  Length 330;
  Best Local Similarity   70.0%;  Pred. No. [illegible];
  Matches    7;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 ILNNQEGWKK 10
           ||:||| | |
Db      27 ILSNQEEWEK 36


RESULT  12
US-08-485-937-12
; Sequence 12, Application US/08485937
; Patent No. 5962633
;  GENERAL INFORMATION:
;    APPLICANT:  Muller, Daniel
;    APPLICANT:  Brownell, Elise
;    APPLICANT:  Delaria, Katherine
;    APPLICANT:  Wallace, Linda
;    TITLE OF INVENTION:  Cytotoxic T-Lymphocyte Antigen as a
;    TITLE OF INVENTION:  Cysteine Protease Inhibitor
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Banner & Allegretti Ltd.
;      STREET:  10 S. Wacker Drive Suite 3000
;      CITY:  Chicago
;      STATE:  IL
;      COUNTRY:  USA
;      ZIP:  60606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC DOS/MS DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/485,937
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/373,[illegible]
;      FILING DATE:  18-MAY-1995
;      APPLICATION NUMBER:  WO PCT/US93/06592
;      FILING DATE:  15-JUL-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/91[illegible]
;      FILING DATE:  17-JUL-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McDonnell, John J
;      REGISTRATION NUMBER:  26,949
;      REFERENCE/DOCKET NUMBER:  9361
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (312) 715-1000
;      TELEFAX:  (312) 715-1234
;  INFORMATION FOR SEQ ID NO:  12:
```

```
;    SEQUENCE CHARACTERISTICS:
;      LENGTH: 96 amino acids
;      TYPE: amino acid
;      STRANDEDNESS: not relevant
;      TOPOLOGY: not relevant
;    MOLECULE TYPE: protein
;    FEATURE:
;      NAME/KEY: Protein
;      LOCATION: 1..96
;      OTHER INFORMATION: /note= "[illegible]"
US-08-485-[illegible]-12

  Query Match             [illegible]
  Best Local Similarity   [illegible]
  Matches   [illegible]  Conservative [illegible]  Mismatches [illegible]

Qy       4 NQEGWKK [illegible]
           [illegible]
Db      26 NEEGWRK [illegible]


RESULT 13
US-08-[illegible]-12
; Sequence 12, Application US/08[illegible]
; Patent No. 597[illegible]
;  GENERAL INFORMATION:
;    APPLICANT:  Muller, [illegible]
;    APPLICANT:  Browse[illegible]
;    APPLICANT:  [illegible], Katherine
;    APPLICANT:  Wallace, [illegible]
;    TITLE OF INVENTION:  Cystatin [illegible]
;    TITLE OF INVENTION:  Cysteine Protease Inhibitors
;    NUMBER OF SEQUENCES:  1[illegible]
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Banner & Allegretti, Ltd
;      STREET:  1[illegible] S. Wacker Drive [illegible]
;      CITY:  Chicago
;      STATE:  IL
;      COUNTRY:  USA
;      ZIP:  60606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.[illegible]
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/[illegible]
;      FILING DATE:  18-MAY-1995
;      CLASSIFICATION:  5[illegible]
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/[illegible]
;      FILING DATE:  [illegible]-JUL-199[illegible]
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  [illegible]
;      FILING DATE:  [illegible]-JUL-199[illegible]
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McDonnell, [illegible]
;      REGISTRATION NUMBER:  2[illegible]
;      REFERENCE/DOCKET NUMBER:  [illegible]
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (312) [illegible]
;      TELEFAX:  (312) [illegible]
;  INFORMATION FOR SEQ ID NO: 1[illegible]:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH: 96 amino acids
;      TYPE: amino acid
;      STRANDEDNESS: not relevant
;      TOPOLOGY: not relevant
;    MOLECULE TYPE: protein
;    FEATURE:
;      NAME/KEY: Protein
;      LOCATION: 1..96
```

```
;      OTHER INFORMATION: /note= "[illegible]"
[illegible]

  Query Match             [illegible]
  Best Local Similarity   [illegible]
  Matches   [illegible]  Mismatches [illegible]  Indels [illegible]  Gaps [illegible]

Qy       4 NQEGWKK

Db      [illegible] NEEGWRK [illegible]


RESULT 14
[illegible]
; Sequence [illegible], Application [illegible]
;  GENERAL INFORMATION:
;    APPLICANT:  Muller, [illegible]
;    APPLICANT:  Brownell, [illegible]
;    APPLICANT:  [illegible], Katherine
;    APPLICANT:  Wallace, [illegible]
;    TITLE OF INVENTION:  Cystatin [illegible] Antibodies
;    TITLE OF INVENTION:  [illegible]
;    NUMBER OF SEQUENCES:  [illegible]
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  [illegible]
;      STREET:  [illegible] Wacker [illegible] Suite [illegible]
;      CITY:  [illegible]
;      STATE:  [illegible]
;      COUNTRY:  USA
;      ZIP:  [illegible]
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  [illegible]
;      COMPUTER:  IBM [illegible]
;      OPERATING SYSTEM:  [illegible]/MS-DOS
;      SOFTWARE:  [illegible] #1[illegible] Version #[illegible]
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  [illegible]
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  [illegible]
;      FILING DATE:  [illegible]-JUL-[illegible]
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  M[illegible]
;      REGISTRATION NUMBER:  [illegible]
;      REFERENCE/DOCKET NUMBER:  [illegible]
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  [illegible]
;      TELEFAX:  [illegible]
;  INFORMATION FOR SEQ ID NO: [illegible]:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH: [illegible] amino acids
;      TYPE: amino acid
;      STRANDEDNESS: [illegible]
;      TOPOLOGY: linear
;    MOLECULE TYPE: protein
[illegible]

  Query Match             [illegible]  Score [illegible]  Length 9[illegible]
  Best Local Similarity   [illegible]  Pred. No. [illegible]
  Matches   [illegible]  Mismatches [illegible]  Indels [illegible]  Gaps [illegible]

Qy       4 N[illegible]WK[illegible]

Db      [illegible] N[illegible]WK[illegible]


RESULT [illegible]
[illegible]
; Sequence [illegible], Application [illegible]
; Patent No. [illegible]
;  GENERAL INFORMATION:
;    APPLICANT:  Chinnadurai, Govindaswamy
;    TITLE OF INVENTION:  APOPTOPSIS REGULATING PROTEINS
;    NUMBER OF SEQUENCES:  35
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SUGHRUE, MION, ZINN, MACPEAK & SEAS
;      STREET:  2100 Pennsylvania Avenue, N.W.
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20037
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0  Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER.  US/08/408,095
;      FILING DATE:  21-MAR-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Mack, Susan J.
;      REGISTRATION NUMBER:  30,951
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)293-7060
;      TELEFAX:  (202)293-7860
;  INFORMATION FOR SEQ ID NO:  21:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  228 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-408-095-21


  Query Match             54.8%;  Score 34;  DB 2;  Length 228;
  Best Local Similarity   50.0%;  Pred. No. [illegible];
  Matches    5;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 ILNNQEGWKK 10
           :|:||  |:|
Db      88 MLSNQASWRK 97


Search completed: August 22, 2001, 14:3[illegible]:1[illegible]
Job time: 367 sec
```

```
                         GenCore version 4.5
               Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on:          August 22, 2001, 14:25:29 ; Search time 61.49 Seconds
                                             (without alignments)
                                             12.993 Million cell updates/sec

Title:           US-08-659-254-8
Perfect score:   62
Sequence:        1 ILNNQEGWKKE 11

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters:       219241

imum DB seq length: 0
mum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       PIR_68:*
                1:  pir1:*
                2:  pir2:*
                3:  pir3:*
                4:  pir4:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                                  SUMMARIES
                   %
Result           Query
  No.    Score   Match Length DB  ID                     Description
------------------------------------------------------------------------------
     1      62   100.0    284  2  A55455                 steroidogenic acut
     2      62   100.0    362  2  JC5386                 steroidogenic acut
     3      57    91.9    285  2  I38248                 steroidogenic acut
     4      51    82.3    285  2  JC4315                 steroidogenic acut
     5      43    69.4    157  2  S62571                 probable ubiquitin
     6      42    67.7    783  2  JC5467                 cellulase (EC 3.2.
     7      42    67.7    824  3  JC7532                 cellulase (EC 3.2.
     8      41    66.1    800  2  F64568                 hypothetical prote
     9      40    64.5    620  2  S64304                 hypothetical prote
    10      40    64.5   1511  2  A53151                 pleiotropic drug r
    11      40    64.5   1529  2  S69688                 hypothetical prote
    12      39    62.9    179  2  S36295                 T cell receptor ga
    13      39    62.9    847  1  A53800                 mixed lineage prot
    14      39    62.9   1711  2  T31357                 1,4-beta glucanase
    15      39    62.9   1742  2  T17110                 cellulase (EC 3.2.
    16      38    61.3    119  2  T23228                 hypothetical prote
    17      38    61.3    480  2  E86689                 prophage ps2 prote
    18      38    61.3   1588  1  BVBYA1                 ARO1 protein - yea
    19      37    59.7    346  2  T10528                 probable gibberell
    20      37    59.7    358  2  S31407                 (2'-5')oligo(A) sy
    21      37    59.7    480  2  T34192                 hypothetical prote
    22      37    59.7    690  2  S62728                 transposase - phag
    23      37    59.7    800  2  A29013                 cellulase (EC 3.2.
    24      37    59.7    822  2  JT0651                 cellulase (EC 3.2.
    25      37    59.7    947  2  I49635                 mouse Dhm1 protein
    26      37    59.7   1245  2  T31953                 hypothetical prote
    27      37    59.7   1271  2  D64237                 hypothetical prote
    28      37    59.7   1289  2  B72354                 conserved hypothet
    29      36    58.1    112  2  T06921                 photosystem II pro
    30      36    58.1    171  2  E82293                 lipoprotein signal
    31      36    58.1    189  2  S53299                 ribonuclease (EC 3
    32      36    58.1    194  1  R5PM24                 ribosomal protein
    33      36    58.1    330  2  JE0376                 Grb-2 related adap
    34      36    58.1    353  2  S56262                 hypothetical prote
    35      36    58.1    366  2  S48310                 neurotoxin type F
    36      36    58.1    368  2  T50174                 probable saccharop
    37      36    58.1    393  2  I38344                 diphosphomevalonat
    38      36    58.1    403  2  T44836                 glycosyltransferas
    39      36    58.1    458  2  H71657                 NADH dehydrogenase
    40      36    58.1    467  2  G84545                 bHLH transcription
    41      36    58.1    505  2  S68204                 trans-cinnamate 4
    42      36    58.1    527  2  I49212                 hypothetical prote
    43      36    58.1    780  2  I29700                 hypothetical prote
    44      36    58.1    799  1  BVKWZ1                 zyg11 protein - Ca
    45      36    58.1    877  2  C64809                 ybaR protein - Esc

                                  ALIGNMENTS

RESULT 1
A55455
steroidogenic acute regulatory protein precursor, mitochondrial - mouse
C;Species: Mus musculus (house mouse)
C;Date: 10-Feb-1995 #sequence_revision 10-Feb-1995 #text_change 08-Oct-1999
C;Accession: A55455
R;Clark, B.J.; Wells, J.; King, S.R.; Stocco, D.M.
J. Biol. Chem. 269, 28314-28322, 1994
A;Title: The purification, cloning, and expression of a novel luteinizing hormone-ind
ry protein (StAR).
A;Reference number: A55455; MUID:95[illegible]
A;Accession: A55455
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-284 <CLA>
A;Cross-references: GB:L36062; GB:[illegible]; G[illegible]; NID:g1236242; PIDN:AAB94783.1; P
C;Keywords: mitochondrion; phosphoprotein
F;1-47/Domain: transit peptide (mitochondrion) #status predicted <[illegible]>

  Query Match             100.0%;  Score 62;  DB 2;  Length 284;
  Best Local Similarity   100.0%;  Pred. No. [illegible];
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ILNNQEGWKKE 11
                |||||||||||
Db         88 ILNNQEGWKKE 98

RESULT 2
JC5386
steroidogenic acute regulatory protein - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 04-Jun-1997 #sequence_revision 4 Jul 1997 #text_change 07-May-1999
C;Accession: JC5386
R;Lee, H.; Ahn, R.S.; Kwon, H.; Soh, J.
Biochem. Biophys. Res. Commun. 230, 528-532, 1997
A;Title: Nucleotide sequence of rat steroidogenic acute regulatory protein complement
A;Reference number: JC5386; MUID:9716[illegible]
A;Accession: JC5386
A;Molecule type: mRNA
A;Residues: 1-362 <LEE>
A;Cross-references: GB:U76419
C;Comment: This protein is a key regulator of the steroidogenesis in acute response t

  Query Match             100.0%;  Score 62;  DB 2;  Length 362;
  Best Local Similarity   100.0%;  Pred. No. 0.0038;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ILNNQEGWKKE 11
                |||||||||||
Db      166 [illegible]NNQ[illegible]WKK[illegible] 17[illegible]


RESULT    3
I38248
steroidogenic acute regulatory protein - human
C;Species: Homo sapiens (man)
C;Date: 29 May 1998 #sequence revision 29 May 1998 #text_change [illegible] Nov [illegible]
C;Accession: I38248; I38[illegible]
R;Sugawara, T.; [illegible] J.A.; Martin, K. [illegible]
Biochemistry 34, 12[illegible] 1995
A;Title: Structure of the human ster[illegible] (StAR) [illegible] (StAR
A;Reference number: I38248; MUID:9600[illegible]
A;Accession: I38248
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-28[illegible]
A;Cross-references: EMBL:[illegible] NID:[illegible] PIDN:AA[illegible]
R;Sugawara, T.; Holt, J.A.; [illegible] M[illegible] W[illegible]
Proc. Natl. Acad. Sci. U.S.A. [illegible]
A;Title: Human steroidogenic acute regulat[illegible]
A;Reference number: I3889[illegible]; MUID:95[illegible]
A;Accession: I388[illegible]
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-28[illegible]
A;Cross-references: EMBL:[illegible] N[illegible] PIDN:AA[illegible]
C;Genetics:
A;Gene: STAR
A;Cross-references: [illegible] M[illegible]
A;Map position: 8p11 [illegible] 8p[illegible]
A;Introns: [illegible]/1; [illegible]


  Query Match        [illegible]
  Best Local Similarity   [illegible]  Pred. No. [illegible]
  Matches  [illegible]  Conservative [illegible]  Mismatches [illegible]

Qy      1 [illegible]NNQEGWKKE 11
          [illegible]
Db     [illegible] [illegible]NQEGWKKE [illegible]


RESULT    4
[illegible]431[illegible]
[illegible]idogenic acute regulatory protein - [illegible]
[illegible]ecies: Bos primigenius taurus (cattle)
[illegible]Date: 29 Nov [illegible] #sequence revision [illegible]
C;Accession: [illegible]4[illegible]
R;Hartung, [illegible] Rust, W.; [illegible] M[illegible]
Biochem. Biophys. Res. Commun. [illegible]
A;Title: Molecular [illegible]
A;Reference number: [illegible]; MUID:[illegible]
A;Accession: [illegible]4[illegible]
A;Molecule type: mRNA
A;Residues: 1-28[illegible] <HA[illegible]
C;Comment: This protein [illegible] acute [illegible]
C;Genetics:
A;Gene: STAR
F;226-264/Region: meta[illegible]lop[illegible]teinase I [illegible]


  Query Match        [illegible]
  Best Local Similarity   [illegible]  Pred. No. [illegible]
  Matches  [illegible]  Conservative [illegible]  Mismatches [illegible]

Qy      1 EENNQEGWKKE 11
          [illegible]
Db     89 EEKDQEGWKKE [illegible]


RESULT    5
[illegible]
```

```
pr[illegible] (EC [illegible]) [illegible] fission yeast (Schizosaccharo
N;A[illegible]
C;Species: [illegible] pombe
C;Date: [illegible] Mar [illegible] #sequence [illegible] Mar [illegible] #text_change [illegible] Feb [illegible]
C;Accession: [illegible]
R;[illegible] M
submitted [illegible] EMBL [illegible] November [illegible]
A;Reference number: [illegible]
A;Accession: [illegible]
A;Status: preliminary
A;Molecule type: DNA
A;Residues: [illegible]
A;Cross-references: EMBL:[illegible] NID:[illegible] PIDN:AA[illegible]
R;[illegible] M.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.
submitted to the EMBL Data Library, November [illegible]
A;Reference number: [illegible]
A;Accession: [illegible]
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: [illegible]
A;Cross-references: EMBL:[illegible] PIDN:CAA[illegible] GSPDB:GN00066; SPDB:SPAC[illegible]
A;Experimental source: strain [illegible]
R;A[illegible] H[illegible] M.; [illegible] A.M
[illegible] Sci. [illegible]
A;Title: [illegible] required for [illegible]
A;Reference number: [illegible]
A;Accession: [illegible]
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: [illegible]
A;Cross-references: EMBL:[illegible] PIDN:AA[illegible]
C;Genetics:
A;Gene: [illegible]
A;Map position: [illegible]
A;Introns: [illegible]
C;Function:
A;Description: required for efficient recovery from DNA damage or S phase arrest
C;Function:
A;Description: [illegible] chromosome segregation
C;Superfamily: [illegible]
C;Keywords: [illegible]


  Query Match        [illegible]  Length [illegible]
  Best Local Similarity   [illegible]  Pred. No. [illegible]
  Matches  [illegible]  Conservative [illegible]  Mismatches [illegible]  Indels [illegible]  Gaps [illegible]

Qy      [illegible]NN[illegible]WK[illegible]

Db      [illegible]N[illegible]WK[illegible]


RESULT    6
[illegible]
cellulase (EC [illegible]) - [illegible]
N;Alternate names: [illegible]
C;Species: [illegible] sp.
C;Date: [illegible] #sequence revision [illegible] #text_change [illegible]
C;Accession: [illegible]
R;Mi[illegible] M.; [illegible]
[illegible]
A;Title: A [illegible] 4 [illegible] sp. [illegible]
A;Reference number: [illegible] MUID:[illegible]
A;Accession: [illegible]
A;Molecule type: DNA
A;Residues: [illegible] <MI[illegible]
A;Cross-references: [illegible]
A;Experimental source: strain [illegible]
A;Note: [illegible] the complete translation [illegible]
C;Superfamily: [illegible] Clostridium xylanase A amino-terminal
C;Keywords: [illegible]
F;[illegible] Clostridium [illegible] A amino-terminal repeat homology <XA
```

```
  Query Match             67.7%;  Score 42;  DB 2;  Length 78[illegible];
  Best Local Similarity   77.8%;  Pred. No. 21;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 ILNNQEGWK 9
           | ||:||||
Db     201 IPNNEEGWK 209


RESULT   7
JC7532
cellulase (EC 3.2.1.4), alkaline - Bacillus sp. (strain KSM-S237)
N;Alternate names: endo-1,4-beta-glucanase
C;Species: Bacillus sp. (strain KSM-S237)
C;Date: 31-Mar-2001 #sequence_revision 31-Mar-2001 #text_change 31-Mar-2001
C;Accession: JC7532; PC7107
R;Hakamada, Y.; Hatada, Y.; Koike, K.; Yoshimatsu, T.; Kawai, S.; Kobayashi, T.; Ito, S.
Biosci. Biotechnol. Biochem. 64, 2281-2289, 2000
A;Title: Deduced amino acid sequence and possible catalytic residues of a thermostable,
A;Reference number: JC7532; MUID:21036886
A;Accession: JC7532
[illegible]lecule type: DNA
[illegible]sidues: 1-824 <HAK>
[illegible]ross-references: DDBJ:AB018420
A;Experimental source: strain KSM-S237
A;Accession: PC7107
A;Molecule type: protein
A;Residues: 31-50 <HA2>
C;Comment: This enzyme is a thermostable, alkaline enzyme used as effective detergent ad
C;Genetics:
A;Gene: Egl-237
C;Keywords: hydrolase, glycosidase


  Query Match             67.7%;  Score 42;  DB 3;  Length 824;
  Best Local Similarity   77.8%;  Pred. No. [illegible];
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 ILNNQEGWK 9
           : |::||||
Db     201 IPNNEEGWK 209


RESULT   8
F64508
hypothetical protein MJ1672 - Methanococcus jannaschii
C;Species: Methanococcus jannaschii
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 22-Oct-1999
C;Accession: F64508
[illegible]ult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake,
[illegible]ich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.;
[illegible], J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.
Science 273, 1058-1073, 1996
A;Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C
A;Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii
A;Reference number: A64300; MUID:96337999
A;Accession: F64508
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-800 <BUL>
A;Cross-references: GB:U67607; GB:L77117; NID:g1592241; PIDN:AAB9[illegible].1; PID:g[illegible], T
C;Genetics:
A;Map position: REV1656207-1653805


  Query Match             66.1%;  Score 41;  DB 2;  Length 800;
  Best Local Similarity   54.5%;  Pred. No. [illegible];
  Matches    6;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 ILNNQEGWKKE 11
           : || || ||
Db     255 LPNNDNGWNKF 265


RESULT   9
S64304
hypothetical protein YGR013w - yeast (Saccharomyces cerevisiae)
N;Alternate names: hypothetical protein G4034
C;Species: Saccharomyces cerevisiae
C;Date: 17-May-1996 #sequence_revision 17-May-1996 #text_change 29-Oct-1999
C;Accession: S64304
R;Rieger, M.; Mueller-Auer, S.; Brueckner, M.; Schaefer, M.
submitted to the Protein Sequence Database, May 1996
A;Reference number: S64071
A;Accession: S64304
A;Molecule type: DNA
A;Residues: 1-620 <RIE>
A;Cross-references: EMBL:Z72798; NID:g[illegible]; PIDN:CAA97024.[illegible]; PID:g1322976; GSPDB:GN00[illegible]
A;Experimental source: strain S288C
C;Genetics:
A;Gene: MIPS:YGR013w
A;Map position: 7R


  Query Match             64.5%;  Score 40;  DB 2;  Length 620;
  Best Local Similarity   60.0%;  Pred. No. 4[illegible];
  Matches    6;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 LNNQEGWKKE 11
           |::||||| :
Db      13 LSSQEGWKSD 22


RESULT  10
A53151
pleiotropic drug resistance protein PDR5 - yeast (Saccharomyces cerevisiae)
N;Alternate names: drug resistance protein YDR1; protein D3542; protein YOR153w; spor
C;Species: Saccharomyces cerevisiae
C;Date: 07-Jul-1995 #sequence_revision 07-Jul-1995 #text_change 02-Feb-2001
C;Accession: A53151; A49730; B49730; S48224; S55359; S67041; S3470[illegible]
R;Bissinger, P.H.; Kuchler, K.
J. Biol. Chem. 269, 4180-4186, 1994
A;Title: Molecular cloning and expression of the Saccharomyces cerevisiae STS1 gene p
A;Reference number: A53151; MUID:94140[illegible]
A;Accession: A53151
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1511 <BIS>
A;Cross-references: GB:X74113; NID:g[illegible]; PIDN:CAA52[illegible].1; PID:g395259
R;Balzi, E.; Wang, M.; Leterme, S.; Van Dyck, L.; Goffeau, A.
J. Biol. Chem. 269, 2206-2214, 1994
A;Title: PDR5, a novel yeast multidrug resistance conferring transporter controlled b
A;Reference number: A49730; MUID:94124[illegible]
A;Accession: A49730
A;Molecule type: DNA
A;Residues: 1-1511 <BAL>
A;Cross-references: GB:L19922; NID:g410[illegible]; PIDN:AAB[illegible].1; PID:g432501
A;Accession: B49730
A;Molecule type: protein
A;Residues: 2-12 <BAW>
R;Hirata, D.; Yano, K.; Miyahara, K.; Miyakawa, T.
Curr. Genet. 26, 285-294, 1994
A;Title: Saccharomyces cerevisiae YDR1, which encodes a member of the ATP-binding cas
A;Reference number: S48224; MUID:95188[illegible]
A;Accession: S48224
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-1511 <HIR>
R;Miyakawa, T.
submitted to the EMBL Data Library, January 1994
A;Reference number: S55359
A;Accession: S55359
A;Molecule type: DNA
A;Residues: 'M',62-1[illegible],'L',172-189,'[illegible]',[illegible]-21[illegible],'[illegible]',21[illegible]-307,'V',309-339,345-475,'H',47
A;Cross-references: EMBL:D26048; NID:g[illegible]; PIDN:BAA05[illegible].1; PID:d1[illegible]; PID:g132
```

```
R;Bordonne, R.; Casa [illegible] A.; Madr[illegible] A.; Mar[illegible]
submitted to the Protein Sequence Database, [illegible]
A;Reference number: [illegible]
A;Accession: S67[illegible]
A;Molecule type: DNA
A;Residues: 1-1511 [illegible]
A;Cross-references: EMBL:[illegible]; NID:g[illegible]; PIDN:[illegible]
A;Experimental source: strain [illegible]
C;Genetics:
A;Gene: SGD:PDR5; [illegible] Y[illegible]
A;Cross-references: [illegible]; MIPS:Y[illegible]w
A;Map position: 15R
C;Superfamily: unassigned ATP-binding cassette protein [illegible]
C;Keywords: ATP; glycoprotein; nucleotide binding; [illegible] transmembrane protein
F;1[illegible]-386/Domain: ATP-binding cassette homology <AB[illegible]>
F;52[illegible]-536/Domain: transmembrane #status predicted <TM[illegible]>
F;[illegible]-576/Domain: transmembrane #status predicted <TM[illegible]>
F;[illegible]-628/Domain: transmembrane #status predicted <TM[illegible]>
F;[illegible]-649/Domain: transmembrane #status predicted <TM[illegible]>
F;663-679/Domain: transmembrane #status predicted <TM[illegible]>
F;7[illegible]-704/Domain: transmembrane #status predicted <TM[illegible]>
F;888-1087/Domain: ATP-binding cassette homology <AB[illegible]>
F;905-912/Region: nucleotide-binding motif A [illegible]
F;1036-125[illegible]/Domain: transmembrane #status predicted <TM[illegible]>
F;1279-1295/Domain: transmembrane #status predicted <TM[illegible]>
F;1330-134[illegible]/Domain: transmembrane #status predicted <TM[illegible]>
F;1358-1374/Domain: transmembrane #status predicted <TM[illegible]>
F;1389-1405/Domain: transmembrane #status predicted <TM[illegible]>
F;1478-149[illegible]/Domain: transmembrane #status predicted <TM[illegible]>


  Query Match              [illegible]1.5%;  Score [illegible];  DB [illegible];  Length [illegible];
  Best Local Similarity    [illegible]%;  Pred. No. [illegible];
  Matches    [illegible];  Conservative  [illegible];  Mismatches  [illegible];  Indels  [illegible];  Gaps  [illegible];

Qy          1 IENNQEGWKK [illegible]
                |||| :|: |
Db        887 IENNVDGWVK [illegible]


RESULT 11
S6[illegible]688
hypothetical protein YOR[illegible]w - yeast (Saccharomyces cerevisiae)
C;Species: Saccharomyces cerevisiae
C;Date: 22 Aug 1997 #sequence_revision [illegible] Sep [illegible] #text_change [illegible]
[illegible]ession: S69688
[illegible]trich, F.S.
submitted to the EMBL Data Library, [illegible]
A;Description: The sequence of [illegible]
A;Reference number: S[illegible]65
A;Accession: S6968[illegible]
A;Molecule type: DNA
A;Residues: 1-1529 [illegible]
A;Cross-references: EMBL:[illegible]; NID:[illegible]; PIDN:AA[illegible]; [illegible]; MIPS:[illegible]w
C;Genetics:
A;Gene: SGD:PDR15
A;Cross-references: SGD:S[illegible]; MIPS:Y[illegible]w
A;Map position: 4R
C;Superfamily: unassigned ATP-binding cassette protein [illegible]
C;Keywords: ATP; nucleotide binding; P-loop
F;186-396/Domain: ATP-binding cassette homology <AB[illegible]>
F;903-1102/Domain: ATP-binding cassette homology <AB[illegible]>
F;920-927/Region: nucleotide-binding motif A [illegible]


  Query Match              [illegible];  Score [illegible];  DB [illegible];  Length [illegible];
  Best Local Similarity    [illegible]%;  Pred. No. [illegible];
  Matches    [illegible];  Conservative  [illegible];  Mismatches  [illegible];  Indels  [illegible];  Gaps  [illegible];

Qy          1 IENNQEGWKK [illegible]
                || : :|
Db        902 IENNVDGWVE [illegible]
```

```
RESULT [illegible]
[illegible]
[illegible] receptor gamma [illegible] precursor - sheep (fragment)
C;Species: [illegible] (sheep)
C;Date: [illegible] #sequence_revision [illegible] #text_change [illegible]
C;Accession: [illegible]
R;[illegible] W.F. [illegible]
EMBO [illegible]
A;Title: [illegible] development
A;Reference number: [illegible]; MUID:[illegible]
A;Accession: [illegible]
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: [illegible]
A;Cross-references: EMBL:[illegible]; NID:g[illegible]; PIDN:AA[illegible]; PID:g[illegible]
C;Superfamily: [illegible] homology
C;Keywords: [illegible]


  Query Match              [illegible];  Score [illegible];  DB [illegible];  Length [illegible];
  Best Local Similarity    [illegible];  Pred. No. [illegible];
  Matches    [illegible];  Conservative  [illegible];  Mismatches  [illegible];  Indels  [illegible];  Gaps  [illegible];

Qy          [illegible] NN[illegible]GWKK

Db          [illegible] N[illegible]GWKK [illegible]


RESULT [illegible]
A[illegible]
mixed lineage protein kinase [illegible]
N;Alternate names: [illegible] protein kinase SPRK
C;Species: Homo sapiens (man)
C;Date: [illegible] Sep [illegible] #sequence_revision [illegible] #text_change [illegible]
C;Accession: [illegible]
R;Gallo, K.A.; Mark, M.R.; Scadden, [illegible]; Wang, [illegible]; Godowski, [illegible]
J. Biol. Chem. [illegible]
A;Title: Identification and characterization of SPRK, a novel src-homology 3 domain[illegible]
A;Reference number: A[illegible]; MUID:[illegible]
A;Accession: A[illegible]
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-[illegible]
A;Cross-references: GB:[illegible]; NID:g[illegible]; PIDN:AAA[illegible]; PID:g[illegible]
R;Ing, Y.L.; [illegible] W.; [illegible]; Gandhi, N.
Oncogene [illegible]
A;Title: MLK[illegible]
A;Reference number: [illegible]; MUID:[illegible]
A;Accession: [illegible]
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: [illegible]
A;Cross-references: GB:[illegible]; NID:g[illegible]; PIDN:AAA[illegible]; PID:g[illegible]
C;Genetics:
A;Gene: [illegible]
A;Cross-references: [illegible]; MIM:[illegible]
A;Map position: [illegible]
C;Superfamily: mixed lineage protein kinase; protein kinase homology; SH3 homology
C;Keywords: ATP; [illegible] phosphotransferase; serine/threonine-specific protein
F;[illegible]/Domain: SH3 homology <SH3>
F;1[illegible]/Domain: protein kinase homology <KIN>
F;[illegible]/Region: protein kinase ATP-binding [illegible]
F;4[illegible]/Region: [illegible]
F;4[illegible]/Region: [illegible]
F;[illegible]


  Query Match              [illegible];  Score [illegible];  DB [illegible];  Length 847;
  Best Local Similarity    [illegible];  Pred. No. [illegible];
  Matches    [illegible];  Conservative  [illegible];  Mismatches  [illegible];  Indels  [illegible];  Gaps  [illegible];

Qy          [illegible]WKK
```

```
             |||| :|
Db     396 QEGWKRE 402


RESULT  14
T31337
1,4-beta-glucanase (EC 3.2.1.-) - Anaerocellum thermophilum (fragment)
C;Species: Anaerocellum thermophilum
C;Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 02-Sep-2000
C;Accession: T31337
R;Zverlov, V.; Mahr, S.; Riedel, K.; Bronnenmeier, K.
Microbiology 144, 457-465, 1998
A;Title: Properties and gene structure of a bifunctional cellulolytic enzyme (CelA) from
omains.
A;Reference number: Z21003; MUID:98154434
A;Accession: T31337
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1711 <ZVE>
A;Cross-references: EMBL:Z86105; NID:e1071129; PID:e350354; PIDN:CAB06786.1
C;Genetics:
A;Gene: celA
  ywords: glycosidase; hydrolase


  Query Match             62.9%;  Score 39;  DB 2;  Length 1711;
  Best Local Similarity   75.0%;  Pred. No. 1.8e+02;
  Matches    6;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 ILNNQEGW 8
             |:||| ||
Db     478 IVNNQSGW 485


RESULT  15
T17120
cellulase (EC 3.2.1.-) precursor, thermoactive - Caldocellum saccharolyticum
C;Species: Caldocellum saccharolyticum
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 17-Nov-2000
C;Accession: T17120; A43745
R;Te'o, V.S.; Saul, D.J.; Bergquist, P.L.
Appl. Microbiol. Biotechnol. 43, 291-296, 1995
A;Title: CelA, another gene coding for a multidomain cellulase from the extreme thermoph
A;Reference number: Z18698; MUID:95336703
A;Accession: T17120
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1742 <TEO>
A;Cross-references: EMBL:L32742; NID:g537499; PID:g537500; PIDN:AAA91086.1
R;Luethi, E.; Bhana Jasmat, N.; Grayling, R.A.; Love, D.R.; Bergquist, P.L.
  l. Environ. Microbiol. 57, 694-700, 1991
  tle: Cloning, sequence analysis, and expression in Escherichia coli of a gene coding
  eference number: A43745; MUID:91247819
A;Accession: A43745
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1516-1544,'A',1546-1742 <LUE>
A;Cross-references: EMBL:M36063; NID:g144232; PIDN:AAA72860.1; PID:g144233
C;Genetics:
A;Gene: celA
C;Keywords: glycosidase; hydrolase


  Query Match             62.9%;  Score 39;  DB 2;  Length 1742;
  Best Local Similarity   75.0%;  Pred. No. 1.8e+02;
  Matches    6;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 ILNNQEGW 8
             |:||| ||
Db     501 IVNNQSGW 508
```

```
Search completed: August 22, 2001, 14:25:30
Job time: 84 sec
```

```
                     GenCore version 4.5
           Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on:           August 22, 2001, 14:29:28  Search time 64.49 Seconds
                                             (without alignments)
                                             10.631 Million cell updates/sec

Title:            US-08-659-254-7
Perfect score:    53
Sequence:         1 AEHGPICMV 9

Scoring table:    BLOSUM62
                  Gapop 10.0 , Gapext 0.5

Searched:         219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters:       219241

[illegible]mum DB seq length: 0
[illegible]mum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :        PIR_68:*
                  1:  pir1:*
                  2:  pir2:*
                  3:  pir3:*
                  4:  pir4:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                              SUMMARIES
                    %
Result            Query
  No.     Score   Match Length DB  ID                          Description
----------------------------------------------------------------------------------
    1       53    100.0    284  2  A5545                       steroidogenic acut
    2       53    100.0    285  2  I3824                       steroidogenic acut
    3       53    100.0    285  2  JC431                       steroidogenic acut
    4       53    100.0    362  2  JC538                       steroidogenic acut
    5       41     77.4    442  2  E8298                       probable oxidoredu
    6     39.5     74.5    325  1  S1575                       acid phosphatase (
    7     39.5     74.5    327  1  JN078                       acid phosphatase (
    8     39.5     74.5    327  1  A4172                       acid phosphatase (
    9     39.5     74.5    340  1  A4609                       acid phosphatase (
   10       39     73.6    670  2  C49934                      transketolase (EC
   11       37     69.8    395  2  JQ0439                      hypothetical 41.4K
   12     35.5     67.0    490  2  E7078                       hypothetical prote
   13       35     66.0    242  2  C25192                      C reactive protein
   14       35     66.0    242  2  B25192                      C reactive protein
   15       35     66.0    557  2  T39308                      asparagine synthas
   16       35     66.0   2531  2  S1818                       notch protein homo
   17       35     66.0   2531  2  A4601                       Notch-1 protein
   18       35     66.0   2555  2  A4004                       notch protein homo
   19       34     64.2     96  2  E8353                       hypothetical prote
   20       34     64.2    402  2  H82755                      2-octaprenyl-6-met
   21       34     64.2    499  2  F69001                      hypothetical prote
   22       34     64.2    500  2  C69428                      hypothetical prote
   23       34     64.2    549  2  H84304                      hypothetical prote
   24       34     64.2    552  2  T30972                      hypothetical prote
   25       34     64.2    583  2  T3271                       hypothetical prote
   26       34     64.2    645  2  H7078                       hypothetical prote
   27       34     64.2    917  2  I49699                      glycoprotein 1.0
   28       34     64.2   2543  2  T31687                      surface antigen
   29       34     64.2   4391  2  A380                        perlecan precurso
```

```
   30       33     62.3    291  2  C83564                      conserved hypothet
   31       33     62.3    523  2  T37488                      actin-like protein
   32       33     62.3    611  2  T44962                      succinate dehydrog
   33       33     62.3    649  2  S67787                      hypothetical prote
   34       33     62.3    661  2  S50734                      WHI3 protein - yea
   35       33     62.3    670  2  A49934                      transketolase (EC
   36       33     62.3    686  1  A59271                      Ra-reactive factor
   37       32     60.4    140  2  T11185                      MLCL536.13 protein
   38       32     60.4    175  2  H85419                      hypothetical prote
   39       32     60.4    181  2  A39582                      liver regeneration
   40       32     60.4    181  2  A54025                      transcription fact
   41       32     60.4    209  2  S54209                      ZIM1 protein - mai
   42       32     60.4    212  2  S23934                      lipase LipA - Baci
   43       32     60.4    222  2  C34223                      transcription fact
   44       32     60.4    254  2  C75612                      phosphate ABC tran
   45       32     60.4    299  2  A75401                      ribulose phosphate

                              ALIGNMENTS

RESULT    1
A55455
steroidogenic acute regulatory protein precursor, mitochondrial - mouse
C;Species: Mus musculus (house mouse)
C;Date: 10-Feb-1995 #sequence_revision 10-Feb-1995 #text_change 08-Oct-1999
C;Accession: A55455
R;Clark, B.J.; Wells, J.; King, S.R.; Stocco, D.M.
J. Biol. Chem. 269, 28314-28322, 1994
A;Title: The purification, cloning, and expression of a novel luteinizing hormone-ind
ry protein (StAR).
A;Reference number: A55455; MUID:95050616
A;Accession: A55455
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-284 <CLA>
A;Cross-references: GB:L36062; GB:L36061; [illegible]; NID:g1236242; PIDN:AAB94783.1; P
C;Keywords: mitochondrion; phosphoprotein
F;1-47/Domain: transit peptide (mitochondrion) #status predicted <TRP>

  Query Match             100.0%;  Score 53;  DB 2;  Length 284;
  Best Local Similarity   100.0%;  Pred. No. 0.015;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         1 AEHGPICMV 9
             |||||||||
Db       217 AEHGPICMV 225


RESULT    2
I38248
steroidogenic acute regulatory protein - human
C;Species: Homo sapiens (man)
C;Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 05-Nov-1999
C;Accession: I38248; I38896
R;Sugawara, T.; Lin, D.; Holt, J.A.; Martin, K.O.; Javitt, N.B.; Miller, W.L.; Straus
Biochemistry 34, 12506-12512, 1995
A;Title: Structure of the human steroidogenic acute regulatory protein (StAR) gene: S
A;Reference number: I38248; MUID:96038[illegible]
A;Accession: I38248
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-285 <RES>
A;Cross-references: EMBL:U29105; NID:[illegible]; PIDN:AA[illegible]4.1; PID:[illegible] 41-98
R;Sugawara, T.; Holt, J.A.; Driscoll, D.; Strauss III, J.F.; Lin, D.; Miller, W.L.; P
Proc. Natl. Acad. Sci. U.S.A. 92, 4778-4782, 1995
A;Title: Human steroidogenic acute regulatory protein: functional activity in COS-1 c
A;Reference number: I38896; MUID:95281[illegible]
A;Accession: I38896
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-285 <RE2>
```

```
A;Cross-references: EMBL:[illegible] NID:[illegible] PIDN:AA[illegible]
C;Genetics:
A;Gene: STAR
A;Cross-references: GDB:[illegible] OMIM:[illegible]
A;Map position: 8p11 [illegible] 8p[illegible]
A;Introns: 22/1; [illegible]

  Query Match [illegible]
  Best Local Similarity [illegible] Pred. No. [illegible]
  Matches [illegible] Conservative [illegible] Mismatches [illegible] Indels [illegible] Gaps [illegible]

Qy          1 AEHGPTCMV [illegible]
              |||||||||
Db        218 AEHGPTCMV [illegible]


[illegible]L    3
[illegible]
steroidogenic acute regulatory protein [illegible]
C;Species: Bos primigenius taurus (cattle)
C;Date: 29 Nov 199[illegible] #sequence_revision [illegible] #text_change [illegible]
C;Accession: [illegible]
R;Hartung, S.; Rust, W.; [illegible] M.; [illegible]
Biochem. Biophys. Res. Commun. 215, [illegible]
A;Title: Molecular [illegible] and [illegible] expression [illegible]
A;Reference number: [illegible] MUID:[illegible]
A;Accession: [illegible]
A;Molecule type: mRNA
A;Residues: 1-285 [illegible]
C;Comment: This protein is an acute controller [illegible]
C;Genetics:
A;Gene: STAR
F;226-264/Region: metalloproteinase [illegible] tissue inhibitor [illegible]

  Query Match            100.0%;  Score [illegible]; DB [illegible]; Length [illegible];
  Best Local Similarity  [illegible]%;  Pred. No. [illegible];
  Matches [illegible] Conservative [illegible] Mismatches [illegible] Indels [illegible] Gaps [illegible]

Qy          1 AEHGPTCMV [illegible]
              |||||||||
Db        218 AEHGPTCMV [illegible]


[illegible]LT    4
[illegible]
steroidogenic acute regulatory protein - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 04 Jan 199[illegible] #sequence_revision [illegible] #text_change [illegible]
C;Accession: [illegible]
R;Lee, H.; Ahn, R.S.; Kwon, H.[illegible]
Biochem. Biophys. Res. Commun. [illegible]
A;Title: Nucleotide sequence of rat steroidogenic acute regulatory protein complementary
A;Reference number: [illegible] MUID:[illegible]
A;Accession: [illegible]
A;Molecule type: mRNA
A;Residues: 1-362 [illegible]
A;Cross-references: GB:[illegible]
C;Comment: This protein is a key regulator [illegible]

  Query Match [illegible]
  Best Local Similarity [illegible] Pred. No. [illegible]
  Matches [illegible] Conservative [illegible] Mismatches [illegible] Indels [illegible] Gaps [illegible]

Qy          1 AEHGPTCMV [illegible]
              |||||||||
Db        295 AEHGPTCMV [illegible]


RESULT    5
[illegible]
```

```
[illegible] oxidoreductase [illegible] Pseudomonas aeruginosa (strain PAO)
C;Species: Pseudomonas aeruginosa
C;Date: [illegible] #sequence_revision [illegible] #text_change [illegible]
C;Accession: [illegible]
R;[illegible] Hickey, M.[illegible]
[illegible]
[illegible]
Nature [illegible]
A;Title: [illegible] Pseudomonas aeruginosa [illegible] opportunistic pa
A;Reference number: A8[illegible] MUID:[illegible]
A;Accession: [illegible]
A;Status: preliminary
A;Molecule type: DNA
A;Residues: [illegible]
A;Cross-references: GB:AE[illegible] NID:[illegible] PIDN:AA[illegible]
A;Experimental source: strain PAO
C;Genetics:
A;Gene: PA[illegible]
C;Superfamily: [illegible]

  Query Match [illegible] Score 41; [illegible] Length 412;
  Best Local Similarity [illegible] Pred. No. [illegible]
  Matches [illegible] Conservative [illegible] Mismatches [illegible] Indels [illegible] Gaps [illegible]

Qy          1 AEH[illegible]
[illegible]
Db        [illegible] AEH[illegible]


RESULT    6
[illegible]
acid phosphatase (EC [illegible]) ACP5 precursor - human
N;Alternate names: [illegible] acid phosphatase (EAP); tartrate-resistant acid phosphatase
C;Species: Homo sapiens (man)
C;Date: 04 Dec 1992 #sequence_revision [illegible] #text_change 03 Mar 2000
C;Accession: [illegible]; A[illegible]; JN0[illegible]; S[illegible]; A4[illegible]; A[illegible]; B[illegible]; A[illegible]; B4[illegible];
R;Lord, D.K.; Cross, N.C.P.; Bevilacqua, M.A.; Rider, S.H.; Gorman, P.A.; Groves, A.V.
Eur. J. Biochem. 189, [illegible]
A;Title: Type 5 acid phosphatase. Sequence, expression and chromosomal [illegible]
A;Reference number: [illegible] MUID:[illegible]
A;Accession: [illegible]
A;Molecule type: mRNA
A;Residues: [illegible]
A;Cross-references: EMBL:X14618; NID:[illegible]; PIDN:[illegible]; PID:[illegible]
A;Experimental source: [illegible]
R;Ketcham, C.M.; Roberts, R.M.; Simmen, R.C.M.; Nick, H.S.
J. Biol. Chem. 264, [illegible]
A;Title: Molecular [illegible] type 5 [illegible] tartrate-resistant acid ph
A;Reference number: A[illegible] MUID:[illegible]
A;Accession: A[illegible]
A;Molecule type: mRNA
A;Residues: [illegible]
A;Cross-references: GB:[illegible]; NID:[illegible]; PIDN:AAA[illegible]; PID:[illegible]
A;Experimental source: [illegible]
R;[illegible], A.; Kim[illegible]; N.[illegible]
Gene [illegible]
A;Title: [illegible] structure of the genes [illegible] type 5
A;Reference number: JN[illegible] MUID:[illegible]
A;Accession: JN[illegible]
A;Molecule type: DNA
A;Residues: [illegible]
A;Cross-references: GB:[illegible]
R;Hayman, A.R.; Warburton, M.J.; [illegible] Chambers, T.[illegible]
Biochem. J. [illegible]
A;Title: Purification and characterization of [illegible] tartrate-resistant acid phosphatase [illegible]
A;Reference number: [illegible] MUID:[illegible]
A;Accession: [illegible]
A;Molecule type: [illegible]
A;Residues: [illegible]
A;Experimental source: [illegible]
R;[illegible], A.[illegible] M.[illegible] K.W.[illegible]
[illegible]
```

```
A;Title: Heterogeneity of hairy cell tartrate resistant acid phosphatase.
A;Reference number: A49133; MUID:93121516
A;Accession: A49133
A;Molecule type: protein
A;Residues: 22-46,'C',48-49,'X',51-55;183-[illegible] <JA4>
A;Experimental source: spleen
A;Note: sequence modified after extraction from NCBI backbone
R;Stepan, J.J.; Lau, K.H.W.; Mohan, S.; Singer, F.R.; Baylink, D.J.
Biochem. Biophys. Res. Commun. 168, 792-800, 1990
A;Title: Purification and N-terminal amino acid sequence of the tartrate resistant acid
A;Reference number: A34638; MUID:90241255
A;Accession: A34638
A;Molecule type: protein
A;Residues: 22-31 <STE>
A;Experimental source: osteoclastoma
R;Stepan, J.J.; Lau, K.H.W.; Mohan, S.; Kraenzlin, M.; Baylink, D.J.
Biochem. Biophys. Res. Commun. 165, 1027-1034, 1989
A;Title: Purification and N-terminal sequences of two tartrate-resistant acid phosphatase
A;Reference number: A33781; MUID:90121201
A;Accession: B33781
A;Molecule type: protein
A;Residues: 22-33,'X',35-37 <ST3>
 xperimental source: hairy cell leukemia spleen
 mment: This enzyme contains a catalytically active binuclear iron complex at the act
C;Genetics:
A;Gene: GDB:ACP5; T5AP
A;Cross-references: GDB:118747; OMIM:1[illegible]164
A;Map position: 19p13.3-19p13.2
A;Introns: 87/3; 130/[illegible]
C;Function:
A;Description: catalyzes the hydrolysis of a wide range of phosphate esters
C;Superfamily: tartrate resistant acid phosphatase; phosphoesterase core homology
C;Keywords: blocked amino end; glycoprotein; iron; metalloprotein; phosphoric monoester
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-182,183-325/Product: acid phosphatase ACP5 #status predicted <MAT>
F;22-182/Domain: acid phosphatase ACP5 chain B #status predicted <CHB>
F;27-112/Domain: phosphoesterase core homology <PEC>
F;183-325/Domain: acid phosphatase ACP5 chain A #status predicted <CHA>
F;33,71,74,242/Binding site: iron (Asp, Asp, Tyr, His) #status predicted
F;71,110,205,240/Binding site: iron (Asp, Asn, His, His) #status predicted
F;111,214/Active site: His #status predicted
F;116,147/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;161-219/Disulfide bonds: #status predicted
F;183/Modified site: blocked amino end (Thr) (in mature form) (partial) #status experime
F;238/Disulfide bonds: interchain (in dimeric form) #status predicted


  Query Match             74.5%;  Score [illegible];  DB 1;  Length 325;
  Best Local Similarity   80.0%;  Pred. No. [illegible];
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    1;  Gaps    1;

          1 AEHGPT-CMV 9
            |||||| |:|
Db      212 AEHGPTHCLV 221


RESULT    7
JN0787
acid phosphatase (EC 3.1.3.2) 5 precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 24-Feb-1994 #sequence_revision 24-Feb-1994 #text_change [illegible]
C;Accession: JN0787; I55486; I55533
R;Cassady, A.I.; King, A.G.; Cross, N.C.; Hume, D.A.
Gene 130, 201-207, 1993
A;Title: Isolation and characterization of the genes encoding mouse and human type 5 aci
A;Reference number: JN0787; MUID:93366175
A;Accession: JN0787
A;Molecule type: DNA
A;Residues: 1-327 <CAS>
A;Cross-references: GB:M99054; NID:g[illegible]; PIDN:AAA37245.1; PID:g[illegible]
R;Reddy, S.V.; Scarcez, T.; Windle, J.J.; [illegible]
J. Bone Miner. Res. 8, 1263-1270, 199[illegible]
A;Title: Cloning and characterization of the 5'-flanking region of the mouse tartrate re
```

```
A;Reference number: I55486; MUID:94078818
A;Accession: I55486
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-48 <RED>
A;Cross-references: GB:M85212; NID:g[illegible]; PIDN:AAA40[illegible].1; PID:g55[illegible]
R;Lacey, D.L.; Erdmann, J.M.; Tan, H.L.
J. Cell. Biochem. 54, 365-371, 1994
A;Title: Interleukin 4 increases type 5 acid phosphatase mRNA expression in murine bo
A;Reference number: I55533; MUID:9425[illegible]
A;Accession: I55533
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 'N',85-1[illegible] <LAC>
A;Cross-references: GB:S70805; NID:g[illegible]; PIDN:AAB[illegible].1; PID:g546968
C;Genetics:
A;Gene: T5AP; TRAP
A;Introns: 89/3; 132/[illegible]; 247/3
C;Function:
A;Description: catalyzes the hydrolysis of a wide range of phosphate esters
A;Note: also functions as a carrier of iron from the uterine endometrium to the fetus
A;Note: this class of phosphatase has optimal activity at acid pH and is resistant to
C;Superfamily: tartrate resistant acid phosphatase; phosphoesterase core homology
C;Keywords: blocked amino end; disulfide bond; glycoprotein; iron; metalloprotein; ph
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-184,185-327/Product: acid phosphatase 5 #status predicted <MAT>
F;22-184/Domain: acid phosphatase 5 chain B #status predicted <CHB>
F;29-114/Domain: phosphoesterase core homology <PEC>
F;85-89/[illegible]/lysosomal sorting sequence
F;185-327/Domain: acid phosphatase 5 chain A #status predicted <CHA>
F;35,73,76,244/Binding site: iron (Asp, Asp, Tyr, His) #status predicted
F;73,112,207,242/Binding site: iron (Asp, Asn, His, His) #status predicted
F;113,216/Active site: His #status predicted
F;118,149/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;163-221/Disulfide bonds: #status predicted
F;185/Modified site: blocked amino end (Thr) (in mature form) (partial) #status exper
F;240/Disulfide bonds: interchain (in dimeric form) #status predicted


  Query Match             74.5%;  Score [illegible];  DB 1;  Length 327;
  Best Local Similarity   80.0%;  Pred. No. [illegible];
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    1;  Gaps    1;

Qy        1 AEHGPT-CMV 9
            |||||| |:
Db      214 AEHGPTRCLV 223


RESULT    8
A41720
acid phosphatase (EC 3.1.3.2) 5 precursor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: A41720
R;Ek-Rylander, B.; Bill, P.; Norgard, M.; Nilsson, S.; Andersson, G.
J. Biol. Chem. 266, 24684-24689, 1991
A;Title: Cloning, sequence, and developmental expression of a type 5, tartrate-resist
A;Reference number: A41720; MUID:92105[illegible]
A;Accession: A41720
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-327 <EKA>
A;Cross-references: GB:M76110; NID:g[illegible]; PIDN:AAA[illegible].1; PID:g207544
C;Superfamily: tartrate-resistant acid phosphatase; phosphoesterase core homology
C;Keywords: phosphoric monoester hydrolase
F;29-114/Domain: phosphoesterase core homology <PEC>


  Query Match             74.5%;  Score [illegible];  DB 1;  Length 327;
  Best Local Similarity   80.0%;  Pred. No. [illegible];
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    1;  Gaps    1;

Qy        1 AEHGPT-CMV 9
```

```
              |:|||:
Db      214 AEHGPIRV.V 223


RESULT   9
A46096
acid phosphatase (EC 3.1.3.2) 5 precursor - pig
N;Alternate names: uteroferrin precursor
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 10 Sep-1999 #sequence revision [illegible] Sep [illegible] #text_change [illegible]
C;Accession: A46096; A33[illegible]; A40928; B40928; A27[illegible]
R;Ling, P.; Roberts, R.M.
J. Biol. Chem. 268, [illegible], 1993
A;Title: Uteroferrin and intracellular tartrate-resistant acid phosphatases are the prod
A;Reference number: A46096; MUID:[illegible]
A;Accession: A46096
A;Status: preliminary; translation not shown
A;Molecule type: mRNA
A;Residues: 1-340 <LIN>
A;Cross-references: GB:M[illegible]; NID:[illegible]; PIDN:AAA[illegible]
R;Simmen, R.C.M.; Srinivas, V.; Roberts, R.M.
DNA 8, 543-554, 1989
A;Title: cDNA sequence, gene organization, and promoter [illegible] mRNA [illegible]
A;Reference number: A33[illegible]; MUID:[illegible]
A;Accession: A33[illegible]
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-182,18[illegible]-340 <SIM>
A;Cross-references: GB:M3[illegible]; NID:[illegible]; PIDN:AAA[illegible]
A;Note: the authors translated the codon [illegible] residue
R;Simmen, R.C.M.; Baumbach, G.A.; Roberts, R.M.
Mol. Endocrinol. [illegible], 1988
A;Title: Molecular cloning and temporal expression [illegible]
A;Reference number: A40928; MUID:[illegible]
A;Accession: A40928
A;Molecule type: mRNA
A;Residues: 186-21[illegible] <[illegible]>
A;Cross-references: GB:M3[illegible]; GB:M[illegible]
A;Accession: B40928
A;Molecule type: mRNA
A;Residues: 2[illegible]2-2[illegible]
A;Cross-references: GB:M3[illegible]
A;Note: the authors translated the [illegible]
R;Hunt, D.F.; Yates, J.R.; [illegible]
Biochem. Biophys. Res. Commun. 144, [illegible], 1987
A;Title: Sequence homology in the metal [illegible]
A;Reference number: A[illegible]; MUID:87[illegible]
A;Accession: A27[illegible]
A;Molecule type: protein
A;Residues: 28-131,'KF[illegible]KE[illegible]
C;Superfamily: tartrate-resistant acid phosphatase; [illegible]
C;Keywords: iron; phosphoric monoester hydrolase
F;35-120/Domain: phosphoesterase [illegible]


  Query Match             [illegible]
  Best Local Similarity   [illegible];  Pred. No. [illegible]
  Matches    8;  Conservative  [illegible];  Mismatches  [illegible];  Indels  [illegible];  Gaps  [illegible]

Qy        1 AEHGPT-MV [illegible]
            |||||| |:|
Db      220 AEHGPTHCIV [illegible]


RESULT  10
C49934
transketolase (EC 2.2.1.1) - Alcaligenes eutrophus [illegible]
C;Species: Alcaligenes eutrophus
C;Date: 19 May 199[illegible] #sequence revision [illegible]
C;Accession: C49934; [illegible]004
R;Schaeferjohann, [illegible]
J. Bacteriol. 175, [illegible], 1993
A;Title: The cbb operons of the facultative [illegible]
A;Reference number: A4[illegible]4; MUID:[illegible]
A;Accession: [illegible]
A;Status: preliminary
A;Molecule type: DNA
A;Residues: [illegible]
A;Cross-references: GB:M[illegible]; NID:g4[illegible]; PIDN:AAA2[illegible]; PID:g4[illegible]
R;Kusemann, [illegible]; Klintworth, [illegible]; Bowien, B.
Gene [illegible]
A;Title: Sequence analysis of the chromosomal and plasmid genes encoding phosph[illegible]
A;Reference number: [illegible]; MUID:[illegible]
A;Accession: [illegible]
A;Molecule type: DNA
A;Residues: [illegible]
A;Experimental source: strain H16
C;Genetics:
A;Genome: plasmid
C;Superfamily: transketolase; [illegible]
C;Keywords: transferase
F;[illegible]/Domain: [illegible]


  Query Match             [illegible]
  Best Local Similarity   [illegible];  Pred. No. [illegible]
  Matches  [illegible];  Conservative  [illegible];  Mismatches  [illegible];  Indels  [illegible];  Gaps  [illegible]

Qy        [illegible] MV [illegible]

Db      [illegible] ENGPHIV [illegible]


RESULT  11
[illegible]
hypothetical 41.4K protein - Streptomyces fradiae transposon Tn45[illegible]
C;Species: Streptomyces fradiae
C;Date: 07 Sep 199[illegible] #sequence revision 07 Sep [illegible] #text_change 21 Jul 2000
C;Accession: [illegible]
R;Siemieniak, [illegible]; Slightom, J.L.; Chung, S.[illegible]
Gene [illegible]
A;Title: Nucleotide sequence of [illegible] Streptomyces fradiae transposable element Tn45[illegible]
A;Reference number: [illegible]; MUID:[illegible]
A;Accession: [illegible]
A;Molecule type: DNA
A;Residues: [illegible]
A;Cross-references: GB:M[illegible]; NID:g15[illegible]; PIDN:AAA88[illegible]; PID:g1[illegible]


  Query Match             [illegible]
  Best Local Similarity   [illegible];  Pred. No. [illegible]
  Matches  [illegible];  Conservative  [illegible];  Mismatches  [illegible];  Indels  [illegible];  Gaps  [illegible]

Qy        [illegible] MV [illegible]

Db      [illegible] V [illegible]


RESULT  12
[illegible]
hypothetical protein [illegible] - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: [illegible] #sequence revision [illegible] #text_change [illegible] Jan 2000
C;Accession: [illegible]
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd
Rajandream, M.A.; Rogers, T.; Rutter, S.; Seeger, K.; Skelton, J.; Squares, S.
Nature [illegible], 1998
A;Authors: Squares, R.; [illegible]; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno
A;Reference number: A7[illegible]; MUID:[illegible]
A;Accession: [illegible]
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: [illegible]
A;Cross-references: [illegible]; NID:[illegible]; PIDN:AA[illegible]
```

```
A;Experimental source: strain H37Rv
C;Genetics:
A;Gene: Rv0888
C;Superfamily: Mycobacterium tuberculosis hypothetical protein Rv0888


  Query Match             67.0%;  Score 3[illegible];  DB 2;  Length 49[illegible];
  Best Local Similarity   53.3%;  Pred. No. 4[illegible];
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    7;  Gaps    [illegible];

Qy        2 EHG-------PTCMV 9
              |||       ||.||
Db      416 EHGPTTPPFAPTCMV 430


RESULT  13
C25192
C-reactive protein chain 1.1 precursor - Atlantic horseshoe crab
C;Species: Limulus polyphemus (Atlantic horseshoe crab)
C;Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change [illegible]
C;Accession: C25192
R;Nguyen, N.Y.; Suzuki, A.; Cheng, S.M.; Zon, G.; Liu, T.
[illegible]Biol. Chem. 261, 10450-10455, 1986
[illegible]tle: Isolation and characterization of Limulus C reactive protein genes.
A;Reference number: A92551; MUID:86278105
A;Accession: C25192
A;Molecule type: DNA
A;Residues: 1-242 <NGU>
A;Cross-references: GB:M14026; NID:g156635; PIDN:AAA28270.1; PID:g156636
C;Genetics:
A;Start codon: ATG
C;Superfamily: C-reactive protein


  Query Match             66.0%;  Score 35;  DB 2;  Length 242;
  Best Local Similarity   100.0%;  Pred. No. [illegible];
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        3 HGPTC 7
              |||||
Db        5 HGPTC 9


RESULT  14
B25192
C-reactive protein chain 3.3 precursor - Atlantic horseshoe crab
C;Species: Limulus polyphemus (Atlantic horseshoe crab)
C;Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change [illegible]
C;Accession: B25192
R;Nguyen, N.Y.; Suzuki, A.; Cheng, S.M.; Zon, G.; Liu, T.
[illegible]Biol. Chem. 261, 10450-10455, 1986
[illegible]tle: Isolation and characterization of Limulus C-reactive protein genes.
[illegible]eference number: A92551; MUID:86278105
A;Accession: B25192
A;Molecule type: DNA
A;Residues: 1-242 <NGU>
A;Cross-references: GB:M14025; NID:g156633; PIDN:AAA28269.1; PID:g156634
C;Genetics:
A;Start codon: ATG
C;Superfamily: C reactive protein


  Query Match             66.0%;  Score 35;  DB 2;  Length 242;
  Best Local Similarity   100.0%;  Pred. No. [illegible];
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        3 HGPTC 7
              |||||
Db        5 HGPTC 9


RESULT  15
T39308
asparagine synthase (glutamine-hydrolyzing) (EC 6.3.5.4) SPBC119.10 [similarity] - fi
C;Species: Schizosaccharomyces pombe
C;Date: 03 Dec 1999 #sequence_revision [illegible] #text_change 21-Jul-2000
C;Accession: T39308; T42010
R;Wood, V.; Rajandream, M.A.; Barrell, B.G.; Skelton, J.; Churcher, C.M.
submitted to the EMBL Data Library, March [illegible]
A;Reference number: Z21843
A;Accession: T39308
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-557 <WOO>
A;Cross-references: EMBL:AL022117; PIDN:CAA[illegible]; GSPDB:GN00067; SPDB:SPBC119.10
A;Experimental source: strain 972h-; cosmid c119
R;Yoshioka, S.; Kato, K.; Nakai, K.; Okayama, H.; Nojima, H.
DNA Res. 4, 363-369, 1997
A;Title: Identification of open reading frames in Schizosaccharomyces pombe cDNAs.
A;Reference number: Z17323; MUID:98172[illegible]
A;Accession: T42010
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: [illegible]ISV
A;Cross-references: EMBL:D89101; NID:g[illegible]; PIDN:BAA13764.1; PID:g17494[illegible]0
A;Experimental source: strain PR745
A;Note: an incorrect initiation codon was used
C;Genetics:
A;Gene: SPDB:SPBC119.10
A;Map position: 2
C;Function:
A;Description: catalyzes the ATP-dependent transfer of the amide group of glutamine t
A;Pathway: asparagine biosynthesis
C;Superfamily: asparagine synthase (glutamine-hydrolyzing)
C;Keywords: asparagine biosynthesis; ligase
F;2-557/Product: asparagine synthase (glutamine-hydrolyzing) #status predicted <MAT>
F;2/Active site: Cys #status predicted


  Query Match             66.0%;  Score 35;  DB 2;  Length 557;
  Best Local Similarity   83.3%;  Pred. No. [illegible];
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        2 EHGPTC 7
              |||| |
Db      109 EHGPAC 114


Search completed: August 22, 2001, 14:2[illegible]
Job time: 33 sec
```

```
                          GenCore version 4.5
              Copyright (c) 1993 - 2001 Compugen Ltd.

OM protein - protein search, using sw model

Run on:          August 22, 2001, 14:30:57 ; Search time 25.8 Seconds
                                            (without alignments)
                                            8.612 Million cell updates/sec

Title:           US-08-659-254-7
Perfect score:   53
Sequence:        1 AEHGPTCMV 9

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters:      93435

   mum DB seq length: 0
   mum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       SwissProt_39:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                                  SUMMARIES
                   %
Result           Query
  No.    Score   Match Length DB  ID                      Description
--------------------------------------------------------------------------
    1       53   100.0    284  1  STAR_MOUSE              P51557 mus musculu
    2       53   100.0    284  1  STAR_RAT                P97826 rattus norv
    3       53   100.0    285  1  STAR_BOVIN              Q28918 bos taurus
    4       53   100.0    285  1  STAR_HORSE              O46689 equus cabal
    5       53   100.0    285  1  STAR_HUMAN              P49675 homo sapien
    6       53   100.0    285  1  STAR_PIG                Q28996 sus scrofa
    7       45    84.9    284  1  STAR_MESAU              P70114 mesocricetu
    8     39.5    74.5    323  1  PPA5_HUMAN              P13686 homo sapien
    9     39.5    74.5    327  1  PPA5_MOUSE              Q05117 mus musculu
   10     39.5    74.5    327  1  PPA5_RAT                P29288 rattus norv
   11     39.5    74.5    340  1  PPA5_PIG                P09889 sus scrofa
   12       39    73.6    670  1  TKTP_ALCEU              P21726 alcaligenes
   13       37    69.8    395  1  YT44_STRLI              P20186 streptomyce
   14     35.5    67.0    490  1  Y888_MYCTU              Q10549 mycobacteri
   15       35    66.0    242  1  CRP1_LIMPO              P06205 limulus pol
   16       35    66.0    242  1  CRP3_LIMPO              P06207 limulus pol
   17       35    66.0   2444  1  NTC1_HUMAN              P46531 homo sapien
   18       35    66.0   2531  1  NTC1_MOUSE              Q01705 mus musculu
   19       35    66.0   2531  1  NTC1_RAT                Q07008 rattus norv
   20       34    64.2    645  1  YL91_MYCTU              Q10384 mycobacteri
   21       34    64.2    917  1  IL6B_MOUSE              Q00560 mus musculu
   22       34    64.2   4393  1  PGBM_HUMAN              P98160 homo sapien
   23       33    62.3    661  1  WHI3_YEAST              P34761 saccharomyc
   24       33    62.3    670  1  TKTC_ALCEU              P21725 alcaligenes
   25       33    62.3    686  1  MAS2_HUMAN              O00187 homo sapien
   26       32    60.4    181  1  ATF3_HUMAN              P18847 homo sapien
   27       32    60.4    181  1  ATF3_MOUSE              Q60765 mus musculu
   28       32    60.4    181  1  ATF3_RAT                P29596 rattus norv
   29       32    60.4    212  1  LIP_BACSU               P37957 bacillus su
   30       32    60.4    216  1  EPA1_XENLA              P52794 xenopus lae
   31       32    60.4    326  1  FRA2_HUMAN              P15408 homo sapien
   32       32    60.4    326  1  FRA2_MOUSE              P47930 mus musculu
   33       32    60.4    327  1  FRA2_RAT                P51145 rattus norv
   34       32    60.4    433  1  HISX_ECOLI              P06988 escherichia
   35       32    60.4    433  1  HISX_SALTY              P10370 salmonella
   36       32    60.4    469  1  HISX_BRAOC              P24226 brassica ol
   37       32    60.4    531  1  GNT3_HUMAN              Q09327 h beta 1,4-
   38       32    60.4    536  1  GNT3_MOUSE              Q10470 m beta 1,4-
   39       32    60.4    536  1  GNT3_RAT                Q02527 r beta 1,4-
   40       32    60.4    908  1  ATMB_SALTY              P22036 salmonella
   41       32    60.4    951  1  AHM1_ARATH              [illegible] arabidopsis
   42       32    60.4   1235  1  CYA4_TRYBB              Q26721 trypanosoma
   43       32    60.4   1108  1  CERB_DROME              P18168 drosophila
   44       32    60.4   2127  1  RRPL_RABVS              P16289 rabies viru
   45       32    60.4   2142  1  RRPL_RABVP              [illegible] rabies viru

                                 ALIGNMENTS

RESULT    1
STAR_MOUSE
ID   STAR_MOUSE      STANDARD;      PRT;   284 AA.
AC   P51557;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN PRECURSOR (LUTEINIZING HORMONE
DE   INDUCED PROTEIN).
GN   STAR.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX   MEDLINE 95050616; PubMed 7961770;
RA   Clark B.J., Wells J., King S.R., Stocco D.M.;
RT   "The purification, cloning, and expression of a novel luteinizing
RT   hormone-induced mitochondrial protein in MA-10 mouse Leydig tumor
RT   cells. Characterization of the steroidogenic acute regulatory protein
RT   (StAR).";
RL   J. Biol. Chem. 269:28314-28322(1994).
CC    !  FUNCTION: PLAYS A KEY ROLE IN STEROID HORMONE SYNTHESIS BY
CC       ENHANCING THE METABOLISM OF CHOLESTEROL INTO PREGNENOLONE
CC       TRANSPORTER THAT BINDS TO AND TRANSPORT CHOLESTEROL THROUGH THE
CC       INTERMEMBRANE SPACE OF THE MITOCHONDRION (PROBABLE).
CC    !  SUBCELLULAR LOCATION: MITOCHONDRIAL.
CC    !  SIMILARITY: BELONGS TO THE STAR FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch)
CC   --------------------------------------------------------------------------
DR   EMBL; L36062; AAB94783.1; -.
DR   MGD; MGI:102760; Star.
DR   InterPro; IPR000799; -.
DR   InterPro; IPR002913; -.
DR   Pfam; PF01852; START; 1.
DR   PRINTS; PR00978; STARPROTEIN.
KW   Lipid-binding; Lipid transport; Transport; Steroidogenesis;
KW   Mitochondrion; Transit peptide.
FT   TRANSIT       1      ?       MITOCHONDRION (POTENTIAL).
FT   CHAIN         ?    284       STEROIDOGENIC ACUTE REGULATORY PROTEIN.
FT   DOMAIN      170    174       POLY-ALA.
SQ   SEQUENCE   284 AA;  31626 MW;  [illegible] CRC64;

  Query Match             100.0%;  Score 53;  DB 1;  Length 284;
  Best Local Similarity   100.0%;  Pred. No. [illegible];
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          1 AEHGPICMV 9
              |||||||||
Db        217 AEHGPICMV 225


RESULT 2
STAR_RAT
ID   STAR_RAT       STANDARD;      PRT;   284 AA.
AC   P97826; P97628;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN PRECURSOR.
GN   STAR.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=WISTAR; TISSUE=Ovary;
RX   MEDLINE=97466080; PubMed=[illegible];
RA   Mizutani T., Sonoda Y., Minegishi T., Wakabayashi K., Miyamoto K.;
RT   "Molecular cloning, characterization and [illegible]
RT   steroidogenic acute regulatory protein (StAR) [illegible]";
RL   Life Sci. 61:1497-1506(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY; TISSUE=Testis;
RX   MEDLINE=97167737; PubMed=9015[illegible];
RA   Lee H.K., Ahn R.[illegible], Kwon H.B., [illegible];
RT   "Nucleotide sequence of rat steroidogenic acute regulatory [illegible]
RT   complementary DNA.";
RL   Biochem. Biophys. Res. Commun. [illegible](1997).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY; TISSUE=Adrenal gland;
RX   MEDLINE=98184087; PubMed=951[illegible];
RA   Ariyoshi N., Kim Y.C., Artemenko I., Bhattacharyya K.K.,
RA   Jefcoate C.R.;
RT   "Characterization of the rat Star gene that encodes the [illegible]
RT   3.5 kilobase pair mRNA. ACTH stimulation [illegible]
RT   precedes elevation of Star mRNA and protein.";
RL   J. Biol. Chem. 273:7610-7619(1998).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY; TISSUE=Adrenal gland;
RX   MEDLINE=99048000; PubMed=983[illegible];
RA   Furukawa A., Miyatake A., Ohnishi T., Ichikawa Y.;
RT   "Steroidogenic acute regulatory protein (StAR) transcripts
RT   constitutively expressed in the adult rat central nervous [illegible]
RT   colocalization of StAR, cytochrome P-450scc (CYPXIA1), and 3beta-
RT   hydroxysteroid dehydrogenase in the rat brain.";
RL   J. Neurochem. 71:2231-2238(1998).
CC   -!- FUNCTION: PLAYS A KEY ROLE IN STEROID HORMONE SYNTHESIS BY
CC       ENHANCING THE METABOLISM OF CHOLESTEROL INTO PREGNENOLONE.
CC       TRANSPORTER THAT BINDS TO AND TRANSPORT CHOLESTEROL THROUGH THE
CC       INTERMEMBRANE SPACE OF THE MITOCHONDRION (PROBABLE).
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL.
CC   -!- SIMILARITY: BELONGS TO THE STAR FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AB001349; BAA19[illegible].1; -.
DR   EMBL; U70419; AAB[illegible].1; -.
DR   EMBL; AF044084; AA[illegible].1; -.
DR   EMBL; AB006007; BAA34[illegible].1; -.
```

```
DR   InterPro; IPR000799; [illegible].
DR   InterPro; IPR002913; START.
DR   Pfam; PF01852; START; 1.
DR   PRINTS; PR00978; STARPROTEIN.
KW   Lipid-binding; Lipid transport; Transport; Steroidogenesis;
KW   Mitochondrion; Transit peptide.
FT   TRANSIT       1     [illegible]       MITOCHONDRION (POTENTIAL).
FT   CHAIN         [illegible]   284       STEROIDOGENIC ACUTE REGULATORY PROTEIN.
FT   CONFLICT      [illegible]   [illegible]       MLLATFK -> MTWESTAVELLHYVELRREEIGMLTVSGFIN
FT                                     EVYEGLLGHTTLIMGHSQNEHSGGVVVNHYLAELFEAHP
FT                                     VYPESVLAYEHE (IN REF. 2).
SQ   SEQUENCE   284 AA;  [illegible] MW;  [illegible] CRC64;

  Query Match             [illegible];  Score [illegible];  DB 1;  Length 284;
  Best Local Similarity   [illegible];  Pred. No. [illegible];
  Matches    [illegible];  Conservative    [illegible];  Mismatches    [illegible];  Indels    [illegible];  Gaps    [illegible];

Qy          1 AEHGPICMV 9
              |||||||||
Db        217 AEHGPICMV 225


RESULT 3
STAR_BOVIN
ID   STAR_BOVIN     STANDARD;      PRT;   285 AA.
AC   [illegible];
DT   01-NOV-1997 (Rel. 35, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN PRECURSOR.
GN   STAR.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Corpus luteum;
RX   MEDLINE=96[illegible]; PubMed=7488004;
RA   Hartung S., Rust W., Balvers M., Ivell R.;
RT   "Molecular cloning and in vivo expression of the bovine steroidogenic
RT   acute regulatory protein.";
RL   Biochem. Biophys. Res. Commun. 215:646-653(1995).
RN   [2]
RP   REVISIONS TO [illegible] AND [illegible].
RA   Hartung S., Rust W., Balvers M., Ivell R.;
RL   Submitted (MAY-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: PLAYS A KEY ROLE IN STEROID HORMONE SYNTHESIS BY
CC       ENHANCING THE METABOLISM OF CHOLESTEROL INTO PREGNENOLONE.
CC       TRANSPORTER THAT BINDS TO AND TRANSPORT CHOLESTEROL THROUGH THE
CC       INTERMEMBRANE SPACE OF THE MITOCHONDRION (PROBABLE).
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL.
CC   -!- TISSUE SPECIFICITY: [illegible] LUTEUM AND ADRENAL GLAND.
CC   -!- SIMILARITY: BELONGS TO THE STAR FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; [illegible]; CAA[illegible].1; -.
DR   EMBL; Y17259; CAA[illegible].1; -.
DR   InterPro; IPR000799; [illegible].
DR   InterPro; IPR002913; [illegible].
DR   Pfam; PF01852; START; 1.
DR   PRINTS; PR00978; STARPROTEIN.
KW   Lipid-binding; Lipid transport; Transport; Steroidogenesis;
KW   Mitochondrion; Transit peptide.
```

```
FT   TRANSIT       1       ?       MITOCHONDRION (POTENTIAL).
FT   CHAIN         ?     285       STEROIDOGENIC ACUTE REGULATORY PROTEIN.
FT   CONFLICT     75      75       H -> Y (IN GENOMIC SEQUENCE).
FT   CONFLICT     99      99       E -> K (IN GENOMIC SEQUENCE).
SQ   SEQUENCE   285 AA;  31845 MW;  EAF05C14EE95E235 CRC64;

  Query Match             100.0%;  Score 5;  DB 1;  Length 285;
  Best Local Similarity   100.0%;  Pred. No. 0.0036;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 AEHGPTCMV 9
            |||||||||
Db      218 AEHGPTCMV 226


RESULT    4
STAR_HORSE
ID   STAR_HORSE     STANDARD;      PRT;   285 AA.
AC   O46689;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
     30-MAY-2000 (Rel. 39, Last annotation update)
     STEROIDOGENIC ACUTE REGULATORY PROTEIN PRECURSOR.
GN   STAR.
OS   Equus caballus (Horse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX   NCBI_TaxID=9796;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Follicle;
RX   MEDLINE=99124482; PubMed=9927292;
RA   Kerban A., Boerboom D., Sirois J.;
RT   "Human chorionic gonadotropin induces an inverse regulation of
RT   steroidogenic acute regulatory protein messenger ribonucleic acid in
RT   theca interna and granulosa cells of equine preovulatory follicles.";
RL   Endocrinology 140:667-674(1999).
CC   -!- FUNCTION: PLAYS A KEY ROLE IN STEROID HORMONE SYNTHESIS BY
CC       ENHANCING THE METABOLISM OF CHOLESTEROL INTO PREGNENOLONE.
CC       TRANSPORTER THAT BINDS TO AND TRANSPORT CHOLESTEROL THROUGH THE
CC       INTERMEMBRANE SPACE OF THE MITOCHONDRION (PROBABLE).
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL.
CC   -!- SIMILARITY: BELONGS TO THE STAR FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
     entities requires a license agreement (See http://www.isb-sib.ch/announce/
     or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AF031697; AAC04704.1; -.
DR   EMBL; AF031696; AAC04703.1; -.
DR   InterPro; IPR000799; -.
DR   InterPro; IPR002913; -.
DR   Pfam; PF01852; START; 1.
DR   PRINTS; PR00978; STARPROTEIN.
KW   Lipid-binding; Lipid transport; Transport; Steroidogenesis;
KW   Mitochondrion; Transit peptide.
FT   TRANSIT       1       ?       MITOCHONDRION (POTENTIAL).
FT   CHAIN         ?     285       STEROIDOGENIC ACUTE REGULATORY PROTEIN.
SQ   SEQUENCE   285 AA;  31852 MW;  B036CEF3299727D CRC64;

  Query Match             100.0%;  Score 5;  DB 1;  Length 285;
  Best Local Similarity   100.0%;  Pred. No. 0.0036;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 AEHGPTCMV 9
            |||||||||
Db      218 AEHGPTCMV 226
```

```
RESULT    5
STAR_HUMAN
ID   STAR_HUMAN     STANDARD;      PRT;   285 AA.
AC   P49675; Q16396;
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN PRECURSOR.
GN   STAR.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Adrenal cortex;
RX   MEDLINE=95281540; PubMed=7761403;
RA   Sugawara T., Holt J.A., Driscoll D., Strauss J.F. III, Lin D.,
RA   Miller W.L., Patterson D., Clancy K.P., Hart I.M., Clark B.J.,
RA   Stocco D.M.;
RT   "Human steroidogenic acute regulatory protein: functional activity in
RT   COS-1 cells, tissue-specific expression, and mapping of the
RT   structural gene to 8p11.2 and a pseudogene to chromosome 13.";
RL   Proc. Natl. Acad. Sci. U.S.A. 92:4778-4782(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Placenta;
RX   MEDLINE=96038208; PubMed=7547998;
RA   Sugawara T., Lin D., Holt J.A., Martin K.O., Javitt N.B., Miller W.L.,
RA   Strauss J.F. III;
RT   "Structure of the human steroidogenic acute regulatory protein (StAR)
RT   gene: StAR stimulates mitochondrial cholesterol 27-hydroxylase
RT   activity.";
RL   Biochemistry 34:12506-12512(1995).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96008588; PubMed=7548191;
RA   Gradi A., Tang Wai R., McBride H.M., Chu L.L., Shore G.C.,
RA   Pelletier J.;
RT   "The human steroidogenic acute regulatory (StAR) gene is expressed in
RT   the urogenital system and encodes a mitochondrial polypeptide.";
RL   Biochim. Biophys. Acta 1258:228-233(1995).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RA   Yu W., Sarginson J., Gibbs R.A.;
RL   Submitted (NOV-1997) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   VARIANT CLAH LEU 182.
RX   MEDLINE=97096837; PubMed=8948562;
RA   Bose H.S., Sugawara T., Strauss J.F. III, Miller W.L.;
RT   "The pathophysiology and genetics of congenital lipoid adrenal
RT   hyperplasia.";
RL   New Engl. J. Med. 335:1870-1878(1996).
RN   [6]
RP   VARIANT ALA 203.
RX   MEDLINE=98112495; PubMed=9452116;
RA   Katsumata N., Tanae A., Shinagawa T., Nagashima-Miyokawa A.,
RA   Shimizu M., Yasunaga T., Tanaka T., Hibi I.;
RT   "A novel frameshift mutation 845delA and a novel polymorphism 852T/A in
RT   the steroidogenic acute regulatory protein gene in a Japanese patient
RT   with congenital lipoid adrenal hyperplasia.";
RL   Hum. Mutat. Suppl. 1:S304-S307(1998).
CC   -!- FUNCTION: PLAYS A KEY ROLE IN STEROID HORMONE SYNTHESIS BY
CC       ENHANCING THE METABOLISM OF CHOLESTEROL INTO PREGNENOLONE.
CC       TRANSPORTER THAT BINDS TO AND TRANSPORT CHOLESTEROL THROUGH THE
CC       INTERMEMBRANE SPACE OF THE MITOCHONDRION (PROBABLE).
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN GONADS, ADRENAL CORTEX AND
CC       KIDNEY.
CC   -!- DISEASE: DEFECTS IN STAR CAUSE CONGENITAL LIPOID ADRENAL
CC       HYPERPLASIA (CLAH). THE MOST SEVERE FORM [illegible]
CC   -!- SIMILARITY: BELONGS TO THE STAR FAMILY.
CC   -
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC
DR   EMBL; U17280; AAC50[illegible]
DR   EMBL; U29105; AA[illegible]50[illegible]
DR   EMBL; U29099; AA[illegible]50[illegible]1.1; JOINED.
DR   EMBL; U29100; AA[illegible]50[illegible]1.1; JOINED.
DR   EMBL; U29101; AA[illegible]50[illegible]1.1; JOINED.
DR   EMBL; U29102; AA[illegible]50[illegible]1.1; JOINED.
DR   EMBL; U29103; AA[illegible]50[illegible]1.1; JOINED.
DR   EMBL; U29104; AA[illegible]50[illegible]1.1; JOINED.
DR   EMBL; S79669; AAB35[illegible].1
DR   EMBL; AF[illegible]; AAB8[illegible]
DR   MIM; 600617;
DR   MIM; 201710;
DR   InterPro; IPR000799; [illegible]
DR   InterPro; IPR00[illegible];
DR   Pfam; PF01852; START; 1.
DR   PRINTS; PR00978; STARPROTEIN.
KW   Lipid-binding; Lipid transport; Transport; Steroidogenesis;
KW   Mitochondrion; Transit peptide; Disease mutation; Polymorphism.
FT   TRANSIT       1     [illegible]      MITOCHONDRION (POTENTIAL).
FT   CHAIN       [illegible]   28[illegible]      STEROIDOGENIC ACUTE REGULATORY PROTEIN.
FT   VARIANT     182    18[illegible]      R -> [illegible] (IN CLAH).
FT                                /FTId=VAR_00562[illegible]
FT   VARIANT     2[illegible]3    2[illegible]      [illegible] -> A
FT                                /FTId=VAR_00[illegible]
SQ   SEQUENCE   285 AA;  [illegible] MW;  [illegible] CRC64;


  Query Match             [illegible];  Score [illegible];  DB [illegible];  Length [illegible];
  Best Local Similarity   [illegible];  Pred. No. [illegible];
  Matches    9;  Conservative    [illegible];  Mismatches    [illegible];  Indels    [illegible];  Gaps    [illegible];

Qy          1 AEHGPTCMV 9
              |||||||||
Db        218 AEHGPTCMV [illegible]


RESULT 6
STAR_PIG
ID   STAR_PIG       STANDARD;      PRT;   285 AA.
AC   Q28996; Q95[illegible]
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN PRECURSOR (StAR).
GN   STAR.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Ovary;
RX   MEDLINE=97200[illegible]; PubMed=[illegible];
RA   Pilon N., Daneau I., Brisson C., Ethier J.-F., [illegible]
RA   Silversides D.W.;
RT   "Porcine and bovine steroidogenic acute regulatory protein (StAR)
RT   gene expression during gestation.";
RL   Endocrinology 13[illegible]:108[illegible]-1[illegible](199[illegible]).
RN   [2]
RP   SEQUENCE OF 1[illegible] FROM N.A.
RC   TISSUE=Corpus luteum;
RX   MEDLINE=97131[illegible]; PubMed=[illegible];
```

```
RA   [illegible]
RA   [illegible]
RT   "Regulation of [illegible] steroidogenic acute regulatory
RT   protein (StAR) by insulin-like growth factor I: synergism with
RT   follicle-stimulating hormone or protein kinase A agonists.";
RL   Endocrinology [illegible]:4[illegible](199[illegible]).
CC   -!- FUNCTION: PLAYS A KEY ROLE IN STEROID HORMONE SYNTHESIS BY
CC       ENHANCING THE METABOLISM OF CHOLESTEROL INTO PREGNENOLONE.
CC       [illegible] BINDS [illegible] AND TRANSFERS [illegible] CHOLESTEROL THROUGH THE
CC       INTERMEMBRANE SPACE OF THE MITOCHONDRION (PROBABLE).
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN ADULT AND FETAL STEROIDOGENIC
CC       [illegible] INCLUDING ADULT TESTIS AND OVARY AND ADULT ADRENAL
CC       GLANDS AS WELL AS STEROIDOGENIC TISSUES OF PREGNANCY, INCLUDING
CC       DEVELOPING FETAL TESTIS, CORPUS LUTEUM, AND PREGNANCY, BUT NOT IN
CC       THE FETAL OVARY.
CC   -!- SIMILARITY: BELONGS TO THE STAR FAMILY.
CC
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC
DR   EMBL; [illegible]; AAB[illegible]
DR   EMBL; [illegible]; AAB[illegible]
DR   InterPro; IPR00[illegible]
DR   InterPro; [illegible]
DR   Pfam; PF01852; START; 1.
DR   PRINTS; PR00978; STARPROTEIN.
KW   Lipid-binding; Lipid transport; Transport; Steroidogenesis;
KW   Mitochondrion; Transit peptide.
FT   TRANSIT       [illegible]      MITOCHONDRION (POTENTIAL).
FT   CHAIN         [illegible]      STEROIDOGENIC ACUTE REGULATORY PROTEIN.
FT   CONFLICT    [illegible]      [illegible] (IN REF. [illegible]).
SQ   SEQUENCE   28[illegible] AA;  [illegible] MW;  [illegible] CRC64;


  Query Match             [illegible];  Score [illegible];  DB [illegible];  Length 28[illegible];
  Best Local Similarity   [illegible];  Pred. No. [illegible];
  Matches    [illegible];  Conservative    [illegible];  Mismatches    [illegible];  Indels    [illegible];  Gaps    [illegible];

Qy          [illegible] AEHGPTCMV [illegible]
              |||||||||
Db        [illegible] AEHGPTCMV [illegible]


RESULT 7
STAR_MESAU
ID   STAR_MESAU     STANDARD;      PRT;   284 AA.
AC   [illegible]
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   [illegible]-MAY-[illegible] (Rel. [illegible], Last annotation update)
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN PRECURSOR (StAR).
GN   STAR.
OS   Mesocricetus auratus (Golden hamster).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC   Mesocricetus.
OX   NCBI_TaxID=[illegible];
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=[illegible]; PubMed=[illegible];
RA   [illegible]
RA   [illegible]
RT   "Adrenocorticotropin [illegible] regulates the levels of the steroidogenic acute
RT   regulatory (StAR) protein mRNA in hamster adrenals.";
RL   [illegible]
CC   -!- FUNCTION: PLAYS A KEY ROLE IN STEROID HORMONE SYNTHESIS BY
CC       ENHANCING THE METABOLISM OF CHOLESTEROL INTO PREGNENOLONE.
CC       TRANSPORTER THAT BINDS TO AND TRANSPORT CHOLESTEROL THROUGH THE
CC       INTERMEMBRANE SPACE OF THE MITOCHONDRION (PROBABLE).
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIA.
CC   -!- SIMILARITY: BELONGS TO THE STAR FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; U66490; AAB06763.1; -.
DR   InterPro; IPR000799; -.
DR   InterPro; IPR002913; -.
DR   Pfam; PF01852; START; 1.
DR   PRINTS; PR00978; STARPROTEIN.
KW   Lipid-binding; Lipid transport; Transport; Steroidogenesis;
KW   Mitochondrion; Transit peptide.
FT   TRANSIT       1      ?       MITOCHONDRION (POTENTIAL).
FT   CHAIN         ?    284       STEROIDOGENIC ACUTE REGULATORY PROTEIN
     SEQUENCE   284 AA;  31497 MW;  F2[illegible]512B CRC64;

  Query Match             84.9%;  Score 4[illegible];  DB 1;  Length 284;
  Best Local Similarity   88.9%;  Pred. No. [illegible]12;
  Matches    8;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 AEHCPTCMV 9
              || ||||||
Db        217 AEQCPTCMV 225


RESULT   8
PPA5_HUMAN
ID   PPA5_HUMAN     STANDARD;      PRT;   323 AA.
AC   P13686;
DT   01-JAN-1990 (Rel. 13, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   TARTRATE-RESISTANT ACID PHOSPHATASE TYPE 5 PRECURSOR (EC 3.1.3.2) (TR
DE   AP) (TARTRATE-RESISTANT ACID ATPASE) (TRATPASE).
GN   ACP5.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
     TISSUE=Placenta;
     MEDLINE=89079710; PubMed=2909539;
RA   Ketcham C.M., Roberts R.M., Simmen R.C.M., Nick H.S.;
RT   "Molecular cloning of the type 5, iron-containing, tartrate-resistant
RT   acid phosphatase from human placenta.";
RL   J. Biol. Chem. 264:557-563(1989).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Placenta;
RX   MEDLINE=90249371; PubMed=2338077;
RA   Lord D.K., Cross N.C.P., Bevilacqua M.A., Rider S.H., Gorman P.A.,
RA   Groves A.V., Moss D.W., Sheer D., Cox T.M.;
RT   "Type 5 acid phosphatase. Sequence, expression and chromosomal
RT   localization of a differentiation-associated protein of the human
RT   macrophage.";
RL   Eur. J. Biochem. 189:287-293(1990).
RN   [3]
RP   SEQUENCE OF 22-37.
RX   MEDLINE=90121201; PubMed=2610679;
RA   Stepan J.J., Lau K.H.W., Mohan S., Kraenzlin M., Baylink D.J.;
RT   "Purification and N-terminal sequence of two tartrate-resistant acid
RT   phosphatases type-5 from the hairy cell leukemia spleen.";
RL   Biochem. Biophys. Res. Commun. 165:1027-1034(1989).
```

```
RN   [4]
RP   SEQUENCE OF 22-31.
RX   MEDLINE=90241255; PubMed=2334436;
RA   Stepan J.J., Lau K.H.W., Mohan S., Singer F.R., Baylink D.J.;
RT   "Purification and N-terminal amino acid sequence of the tartrate-
RT   resistant acid phosphatase from human osteoclastoma: evidence for a
RT   single structure.";
RL   Biochem. Biophys. Res. Commun. 168:792-800(1990).
RN   [5]
RP   SEQUENCE OF 22-64 AND 181-200.
RX   MEDLINE=8937[illegible]; PubMed=2775236;
RA   Hayman A.R., Warburton M.J., Pringle J.A.S., Coles B., Chambers T.J.;
RT   "Purification and characterization of a tartrate-resistant acid
RT   phosphatase from human osteoclastomas.";
RL   Biochem. J. 261:601-609(1989).
RN   [6]
RP   SUBUNITS.
RX   MEDLINE=9133[illegible]78; PubMed=18727[illegible];
RA   Hayman A.R., Dryden A.J., Chambers T.J., Warburton M.J.;
RT   "Tartrate-resistant acid phosphatase from human osteoclastomas is
RT   translated as a single polypeptide.";
RL   Biochem. J. 277:631-634(1991).
CC   -!- CATALYTIC ACTIVITY: AN ORTHOPHOSPHORIC MONOESTER + H(2)O = AN
CC       ALCOHOL + ORTHOPHOSPHATE.
CC   -!- COFACTOR: BINDS TWO IRON ATOMS.
CC   -!- SUBUNIT: EXISTS EITHER AS MONOMER OR, AFTER PROTEOLYTIC
CC       PROCESSING, AS A DIMER OF TWO CHAINS LINKED BY DISULFIDE BOND(S).
CC   -!- SUBCELLULAR LOCATION: LYSOSOMAL.
CC   -!- DISEASE: THIS RELATIVELY MINOR INTRACELLULAR ISOZYME OF ACID
CC       PHOSPHATASE CAN BECOME THE DOMINANT ISOZYME IN CERTAIN
CC       PATHOLOGICAL STATES (GAUCHER'S AND HODGKIN'S DISEASES, THE HAIRY
CC       CELL, THE B CELL, AND THE T CELL LEUKEMIAS).
CC   -!- SIMILARITY: TO OTHER TARTRATE RESISTANT ACID PHOSPHATASES.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; J04430; AAA76849.1; -.
DR   EMBL; X14618; CAA32771.1; -.
DR   PIR; A31812; A31812.
DR   PIR; A33781; A33781.
DR   PIR; B33781; B33781.
DR   PIR; A34638; A34638.
DR   PIR; S05251; S05251.
DR   PIR; S05255; S05255.
DR   PIR; S15752; S15752.
DR   MIM; 171640; -.
KW   Hydrolase; Glycoprotein; Iron; Lysosome; Signal
FT   SIGNAL        1     21
FT   CHAIN        22    323       TARTRATE-RESISTANT ACID PHOSPHATASE TYPE
FT                                5.
FT   CARBOHYD    116    116       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD    147    147       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CONFLICT     45     46       AR -> [illegible]P (IN REF. [illegible]).
FT   CONFLICT     47     47       E -> [illegible] (IN REF. 5).
FT   CONFLICT     50     50       N -> W (IN REF. 5).
FT   CONFLICT    177    178       LT -> DVKL (IN REF. 2).
SQ   SEQUENCE   323 AA;  36357 MW;  F4[illegible]28D4[illegible] CRC64;

  Query Match             74.5%;  Score 39.5;  DB 1;  Length 323;
  Best Local Similarity   80.0%;  Pred. No. 1.6;
  Matches    8;  Conservative    1;  Mismatches    [illegible];  Indels    1;  Gaps    1;

Qy          1 AEHCPT-CMV 9
              ||||| |:|
Db        210 AEHCPIECLV 219
```

```
RESULT [illegible]
PPA5_MOUSE
ID   PPA5_MOUSE     STANDARD;      PRT;   [illegible] AA.
AC   Q05117;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   TARTRATE-RESISTANT ACID PHOSPHATASE TYPE 5 PRECURSOR [illegible]
DE   AP) (TARTRATE-RESISTANT ACID ATPASE) (TRATPASE).
GN   ACP5 OR TRAP OR TSAP.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; [illegible]
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=[illegible];
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93[illegible]; PubMed=[illegible];
RA   Cassady A.I., King A.G., Cross N.[illegible]
RT   "Isolation and characterization of the [illegible]
RT   type 5 acid phosphatase.";
RL   Gene 130:201-20[illegible](199[illegible]).
RN   [2]
RP   SEQUENCE OF 1-48 FROM N.A.
RC   TISSUE=Spleen;
RX   MEDLINE=94[illegible]; PubMed=[illegible];
RA   Reddy S.V., Scarcez T., Windle J.J., Leach R.J., [illegible]
RA   Chirgwin J.M., Chou J.Y., Roodman G.D.;
RT   "Cloning and characterization of the [illegible]
RT   tartrate resistant acid phosphatase gene.";
RL   J. Bone Miner. Res. 8:1263-1270(1993).
CC   -!- FUNCTION: MAY PLAY A ROLE IN THE PROTEOLYSIS OF BONE [illegible] IN THE
CC       OSTEOCLASTIC TRAP ACTS ON NUCLEOTIDE TRI- AND DIPHOSPHATES WITH
CC       HIGHER AFFINITY, COMPARED WITH OTHER SUBSTRATES.
CC   -!- CATALYTIC ACTIVITY: AN ORTHOPHOSPHORIC MONOESTER + H(2)O = AN
CC       ALCOHOL + ORTHOPHOSPHATE.
CC   -!- COFACTOR: BINDS TWO IRON ATOMS.
CC   -!- SUBUNIT: EXISTS EITHER AS MONOMER OR, AFTER PROTEOLYTIC
CC       PROCESSING, AS A DIMER OF TWO CHAINS LINKED BY DISULFIDE BOND(S).
CC   -!- SUBCELLULAR LOCATION: LYSOSOMAL.
CC   -!- TISSUE SPECIFICITY: CHARACTERISTIC CONSTITUENT [illegible]
CC   -!- SIMILARITY: HIGH IDENTITY TO PORCINE TYPE 5 PROTEIN AND TO
CC       BOVINE SPLEEN PHOSPHOPROTEIN PHOSPHATASE.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; M99054; AAA[illegible]24.1; -.
DR   EMBL; M85212; AAA1047[illegible].1; -.
DR   PIR; JN0787; JN0787.
DR   MGD; MGI:87883; Acp5.
KW   Hydrolase; Glycoprotein; Iron; Lysosome; Signal.
FT   SIGNAL        1     2[illegible]       BY SIMILARITY.
FT   CHAIN        2[illegible]    32[illegible]      TARTRATE-RESISTANT ACID PHOSPHATASE TYPE
FT                                5.
FT   CARBOHYD    11[illegible]   11[illegible]       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD    14[illegible]   14[illegible]       N-LINKED (GLCNAC...) (POTENTIAL).
SQ   SEQUENCE   32[illegible] AA;  [illegible] MW;  [illegible] CRC64;

  Query Match             [illegible]
  Best Local Similarity   [illegible];  Pred. No. [illegible];
  Matches  [illegible];  Conservative  [illegible];  Mismatches  [illegible];  Indels  [illegible];  Gaps  [illegible];

Qy          1 AEHGPI[illegible]MV[illegible]
              ||||||
Db        214 AEHGPIRCLV[illegible]
```

```
RESULT [illegible]
PPA5_RAT
ID   PPA5_RAT       STANDARD;      PRT;   [illegible] AA.
AC   [illegible]
DT   01-DEC-1992 (Rel. 24, Created)
DT   01-DEC-1992 (Rel. 24, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   TARTRATE-RESISTANT ACID PHOSPHATASE TYPE 5 PRECURSOR (EC [illegible]) (TR
DE   AP) (TARTRATE-RESISTANT ACID ATPASE) (TRATPASE).
GN   ACP5.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=[illegible];
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Bone;
RX   MEDLINE=[illegible]; PubMed=[illegible];
RA   Ek-Rylander B., [illegible], Norgard M., [illegible], Andersson G.;
RT   "Cloning, sequence, and developmental expression of a type 5,
RT   tartrate-resistant, acid phosphatase of rat bone.";
RL   [illegible]
RN   [2]
RP   SEQUENCE OF [illegible].
RX   MEDLINE=[illegible]; PubMed=[illegible];
RA   Ek-Rylander B., [illegible], Andersson G.;
RT   "Characterization of a tartrate-resistant acid phosphatase (ATPase)
RT   from rat bone: hydrodynamic properties and N-terminal amino acid
RT   sequence.";
RL   J. Bone Miner. Res. [illegible].
CC   -!- FUNCTION: MAY PLAY A ROLE IN THE [illegible] OF BONE RESORPTION. THE
CC       OSTEOCLASTIC TRAP ACTS ON NUCLEOTIDE TRI- AND DIPHOSPHATES WITH
CC       HIGHER AFFINITY, COMPARED WITH OTHER SUBSTRATES.
CC   -!- CATALYTIC ACTIVITY: AN ORTHOPHOSPHORIC MONOESTER + H(2)O = AN
CC       ALCOHOL + ORTHOPHOSPHATE.
CC   -!- COFACTOR: BINDS TWO IRON ATOMS.
CC   -!- SUBUNIT: EXISTS EITHER AS MONOMER OR, AFTER PROTEOLYTIC
CC       PROCESSING, AS A DIMER OF TWO CHAINS LINKED BY DISULFIDE BOND(S).
CC   -!- SUBCELLULAR LOCATION: LYSOSOMAL.
CC   -!- TISSUE SPECIFICITY: CHARACTERISTIC CONSTITUENT OF OSTEOCLASTS AND
CC       SOME MONONUCLEAR PRECURSOR CELLS; PREFERENTIALLY EXPRESSED IN
CC       SKELETAL TISSUES.
CC   -!- SIMILARITY: [illegible] TARTRATE RESISTANT ACID PHOSPHATASES.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; M[illegible]; AAA4[illegible].
DR   PIR; A41[illegible]; A41[illegible].
KW   Hydrolase; Glycoprotein; Iron; Lysosome; Signal.
FT   SIGNAL        1     2[illegible]
FT   CHAIN        2[illegible]    32[illegible]      TARTRATE-RESISTANT ACID PHOSPHATASE TYPE
FT                                5.
FT   CARBOHYD    12[illegible]   12[illegible]       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD    14[illegible]   14[illegible]       N-LINKED (GLCNAC...) (POTENTIAL).
SQ   SEQUENCE   327 AA;  [illegible] MW;  [illegible] CRC64;

  Query Match             [illegible];  Score [illegible];  DB [illegible];  Length [illegible];
  Best Local Similarity   [illegible];  Pred. No. [illegible];
  Matches  [illegible];  Conservative  [illegible];  Mismatches  [illegible];  Indels  [illegible];  Gaps  [illegible];

Qy          1 AEHGPI[illegible]MV[illegible]
              [illegible]
Db          [illegible] AEHGPIRCLV[illegible]
```

```
RESULT  11
PPA5_PIG
ID   PPA5_PIG       STANDARD;      PRT;   340 AA.
AC   P09889;
DT   01-MAR-1989 (Rel. 10, Created)
DT   01-OCT-1993 (Rel. 27, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   TARTRATE-RESISTANT ACID PHOSPHATASE TYPE 5 PRECURSOR (EC 3.1.3.2) (TR
DE   AP) (TARTRATE-RESISTANT ACID ATPASE) (TRATPASE) (UTEROFERRIN) (UF).
GN   ACP5.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC   TISSUE=Spleen;
RX   MEDLINE=93216620; PubMed 8463220;
RA   Ling P., Roberts R.M.;
RT   "Uteroferrin and intracellular tartrate-resistant acid phosphatases
RT   are the products of the same gene.";
     J. Biol. Chem. 268:6896-6902(1993).
     [2]
     SEQUENCE FROM N.A.
RX   MEDLINE=90091741; PubMed-2598770;
RA   Simmen R.C.M., Srinivas V., Roberts R.M.;
RT   "cDNA sequence, gene organization and progesterone induction of mRNA
RT   for uteroferrin, a porcine uterine iron transport protein.";
RL   DNA 8:543-554(1989).
RN   [3]
RP   SEQUENCE OF 186-215 AND 247-301 FROM N.A.
RX   MEDLINE-88288243; PubMed-2969454;
RA   Simmen R.C.M., Baumbach G.A., Roberts R.M.;
RT   "Molecular cloning and temporal expression during pregnancy of the
RT   messenger ribonucleic acid encoding uteroferrin, a
RT   progesterone-induced uterine secretory protein.";
RL   Mol. Endocrinol. 2:253-262(1988).
RN   [4]
RP   SEQUENCE OF 108-195 FROM N.A.
RX   MEDLINE=96036743; PubMed=7587387;
RA   Yasue H., Kusumoto H., Mikami H.;
RT   "Assignment of the uteroferrin gene (ACP5) to swine chromosome 2q12-
RT   >q21 by fluorescence in situ hybridization.",
RL   Cytogenet. Cell Genet. 71:249-252(1995).
RN   [5]
RP   PARTIAL SEQUENCE.
RX   MEDLINE=87213307; PubMed=3579955;
RA   Hunt D.F., Yates J.R. III, Shabanowitz J., Zhu N.-Z., Zirino T.,
RA   Averill B.A., Daurat-Larroque S.T., Shewale J.G., Roberts R.M.,
RA   Brew K.;
     "Sequence homology in the metalloproteins; purple acid phosphatase
     from beef spleen and uteroferrin from porcine uterus.",
     Biochem. Biophys. Res. Commun. 144:1154-1160(1987).
CC   -!- FUNCTION: UTEROFERRIN IS A PHOSPHOPROTEIN PHOSPHATASE, SYNTHESIZED
CC       IN RESPONSE TO PROGESTERONE. IT APPEARS TO FUNCTION IN
CC       TRANSPLACENTAL TRANSPORT OF IRON IN PIG.
CC   -!- CATALYTIC ACTIVITY: AN ORTHOPHOSPHORIC MONOESTER + H(2)O = AN
CC       ALCOHOL + ORTHOPHOSPHATE.
CC   -!- COFACTOR: BINDS TWO IRON ATOMS.
CC   -!- SUBCELLULAR LOCATION: SECRETED.
CC   -!- SIMILARITY: TO OTHER TARTRATE-RESISTANT ACID PHOSPHATASES.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long as its content is in no way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; M98553; AAA31129.1; -.
DR   EMBL; M30284; AAA31139.1; -.
DR   EMBL; M30283; AAA31138.1; -.
DR   EMBL; M31214; AAA31134.1; -.
DR   EMBL; M31215; AAA31135.1; -.
DR   EMBL; S80099; AAB14321.1; -.
DR   PIR; A27035; A27035.
DR   PIR; A33318; A33318.
KW   Hydrolase; Glycoprotein; Iron; Iron transport; Signal.
FT   SIGNAL        1     20
FT   CHAIN        21    340       TARTRATE-RESISTANT ACID PHOSPHATASE TYPE
FT                                5.
FT   CARBOHYD    124    124       N-LINKED (GLCNAC...).
FT   CARBOHYD    155    155       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CONFLICT    183    184       MISSING (IN REF. 2)
SQ   SEQUENCE   340 AA;  38011 MW;  CFF15[illegible]128[illegible]79F7[illegible]64;

  Query Match             74.5%;  Score [illegible];  DB [illegible];  Length 340;
  Best Local Similarity   80.0%;  Pred. No. [illegible];
  Matches    8;  Conservative    1;  Mismatches    0;  Indels    1;  Gaps    1;

Qy        1 AEHCPI-CMV [illegible]
            |||||| :|
Db      220 AEHCPIHCLV 229


RESULT  12
TKTP_ALCEU
ID   TKTP_ALCEU     STANDARD;      PRT;   670 AA.
AC   P21726;
DT   01-MAY-1991 (Rel. 18, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   TRANSKETOLASE, PLASMID (EC 2.2.1.1) (TK).
GN   CBBTP.
OS   Alcaligenes eutrophus (Ralstonia eutropha).
OG   Plasmid megaplasmid pHG1.
OC   Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
OC   Ralstonia.
OX   NCBI_TaxID=510;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=H16 / ATCC 17699;
RX   MEDLINE 94042908; PubMed 8226684;
RA   Schaeferjohann J., Yoo J.-G., Kusian B., Bowien B.;
RT   "The cbb operons of the facultative chemoautotroph Alcaligenes
RT   eutrophus encode phosphoglycolate phosphatase.";
RL   J. Bacteriol. 175:7329-7340(1993).
RN   [2]
RP   SEQUENCE OF 1-77 FROM N.A.
RC   STRAIN-H16 / ATCC 17699;
RX   MEDLINE-90152372; PubMed-2559876;
RA   Kossmann J., Klintworth R., Bowien B.;
RT   "Sequence analysis of the chromosomal and plasmid genes encoding
RT   phosphoribulokinase from Alcaligenes eutrophus."
RL   Gene 85:247-252(1989).
CC   -!- CATALYTIC ACTIVITY: SEDOHEPTULOSE 7-PHOSPHATE + D-GLYCERALDEHYDE
CC       3-PHOSPHATE = D-RIBOSE 5-PHOSPHATE + D-XYLULOSE 5-PHOSPHATE.
CC   -!- COFACTOR: THIAMINE PYROPHOSPHATE.
CC   -!- PATHWAY: CALVIN CYCLE.
CC   -!- SIMILARITY: BELONGS TO THE TRANSKETOLASE FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long as its content is in no way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; M68905; AAA20194.1; -.
DR   EMBL; M33562; -; NOT_ANNOTATED_CDS.
DR   PIR; PQ0045; PQ0045.
DR   HSSP; P23254; 1AY0.
DR   InterPro; IPR000360; -.
DR   Pfam; PF004[illegible]; [illegible]
DR   PROSITE; PS00[illegible]; TRANSKETOLASE_[illegible]
DR   PROSITE; PS00[illegible]; TRANSKETOLASE_[illegible]
KW   Transferase; Thiamine pyrophosphate; [illegible]
SQ   SEQUENCE   [illegible] AA;  [illegible] MW;  [illegible]

  Query Match              [illegible]%;  Score [illegible];  DB [illegible];  Length [illegible];
  Best Local Similarity    [illegible]%;  Pred. No. [illegible];
  Matches    6;  Conservative  [illegible];  Mismatches  [illegible];  Indels  [illegible];  Gaps  [illegible];

Qy          2 EHGPTCMV [illegible]
              |: |||:
Db        [illegible]16 ENGPTC.V [illegible]


RESULT  13
[illegible]STREF
ID   Y144 STREF     STANDARD;      PRT;   [illegible] AA.
AC   P20186;
DT   01 FEB 1991 (Rel. [illegible], Created)
DT   01 FEB 1991 (Rel. [illegible], Last sequence update)
DT   01 OCT 1996 (Rel. 34, Last annotation update)
DE   HYPOTHETICAL [illegible] KDA PROTEIN IN TRANSPOSON [illegible]
OS   Streptomyces fradiae.
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Streptomycineae; Streptomycetaceae; [illegible]
OX   NCBI_TaxID=1906;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TRANSPOSON Tn4556;
RX   MEDLINE=90185230; PubMed=[illegible];
RA   Siemieniak D.R., Slightom J.L., [illegible]
RT   "Nucleotide sequence of Streptomyces fradiae [illegible]
RT   Tn4556: a class II transposon related to [illegible]"
RL   Gene 86:1-9(1990).
CC   [illegible]
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   [illegible]
DR   EMBL; M29297; AAA885[illegible]; [illegible]
DR   PIR; JQ0430; [illegible]
KW   Hypothetical protein; Transposable element.
SQ   SEQUENCE   [illegible] AA;  [illegible] MW;  [illegible]

  Query Match              [illegible]%;  Score [illegible];  DB [illegible];  Length [illegible];
  Best Local Similarity    [illegible]%;  Pred. No. [illegible];
  Matches    [illegible];  Conservative  [illegible];  Mismatches  [illegible];  Indels  [illegible];  Gaps  [illegible];

Qy          2 EHGPTCMV [illegible]
              :||: |:
Db        [illegible] QHGPPCLV [illegible]


RESULT  14
Y888_MYCTU
ID   Y888_MYCTU     STANDARD;      PRT;   [illegible] AA.
AC   Q10549;
DT   01 OCT 1996 (Rel. 34, Created)
DT   01 OCT 1996 (Rel. 34, Last sequence update)
DT   30 MAY 2000 (Rel. 39, Last annotation update)
DE   HYPOTHETICAL [illegible] KDA PROTEIN RV0888.
GN   RV0888 OR MTCY31.[illegible].
OS   Mycobacterium tuberculosis.
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium; [illegible]
OX   NCBI_TaxID=1773;
```

```
RN   [illegible]
RP   SEQUENCE FROM N.A.
RC   STRAIN=H37RV;
RX   MEDLINE=[illegible]; PubMed=[illegible];
RA   [illegible], Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA   [illegible], Eiglmeier K., Gas S., Barry C.E. III, Tekaia F.,
RA   Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA   [illegible], Devlin K., [illegible], [illegible], Hamlin N., Holroyd S.,
RA   Hornsby T., [illegible], Krogh A., McLean J., Moule S., Murphy L.,
RA   Oliver K., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA   Rutter S., Seeger K., Skelton S., Squares R., Squares S., Sulston J.E.,
RA   Taylor K., Whitehead S., Barrell B.G.;
RT   "Deciphering the biology of Mycobacterium tuberculosis from the
RT   complete genome sequence.";
RL   Nature 393:537-544(1998).
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC   [illegible]
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   [illegible]
DR   EMBL; Z[illegible]; CAA[illegible]; [illegible]
DR   [illegible]
KW   Hypothetical protein; Transmembrane.
FT   TRANSMEM     [illegible]    [illegible]       POTENTIAL.
FT   TRANSMEM     [illegible]    [illegible]       POTENTIAL.
FT   TRANSMEM     [illegible]    [illegible]       POTENTIAL.
FT   TRANSMEM     [illegible]    [illegible]       POTENTIAL.
FT   TRANSMEM     [illegible]    [illegible]       POTENTIAL.
SQ   SEQUENCE   [illegible] AA;  [illegible] MW;  [illegible]

  Query Match              [illegible]%;  Score [illegible];  DB [illegible];  Length [illegible];
  Best Local Similarity    [illegible]%;  Pred. No. [illegible];
  Matches    [illegible];  Conservative  [illegible];  Mismatches  [illegible];  Indels  [illegible];  Gaps  [illegible];

Qy          [illegible] EHG[illegible]  [illegible] MV [illegible]

Db        [illegible] [illegible] MV [illegible]


RESULT  [illegible]
[illegible]LIMPO
ID   [illegible]_LIMPO     STANDARD;      PRT;   [illegible] AA.
AC   [illegible];
DT   [illegible] JAN [illegible] (Rel. [illegible], Created)
DT   [illegible] JAN [illegible] (Rel. [illegible], Last sequence update)
DT   [illegible] NOV [illegible] (Rel. [illegible], Last annotation update)
DE   [illegible]
OS   Limulus polyphemus (Atlantic horseshoe crab).
OC   Eukaryota; Metazoa; Arthropoda; Chelicerata; Merostomata; Xiphosura;
OC   Limulidae; Limulus.
OX   NCBI_TaxID=[illegible];
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=[illegible]; PubMed=[illegible];
RA   Nguyen N.Y., Suzuki A., [illegible], [illegible] Y.;
RT   "[illegible] of a putative protein [illegible].";
RL   [illegible]
RN   [2]
RP   SEQUENCE [illegible] NTS, AND [illegible] LINKAGE [illegible]
RX   MEDLINE=[illegible]; PubMed=[illegible];
RA   Nguyen N.Y., Suzuki A., [illegible];
RT   "The [illegible]
RT   [illegible]";
RL   [illegible]
RN   [3]
RP   [illegible]
RX   MEDLINE=[illegible]; PubMed=[illegible];
```

```
RA   Srinivasan N., White H.E., Emsley J., Wood S.P., Pepys M.B.,
RA   Blundell T.L.;
RT   "Comparative analyses of pentraxins: implications for protomer
RT   assembly and ligand binding.";
RL   Structure 2:1017-1027(1994).
CC   -!- FUNCTION: MIGHT SERVE THE ROLE OF IMMUNOGLOBULINS.
CC   -!- SUBUNIT: HOMOPENTAMER. PENTAXIN (OR PENTRAXIN) HAVE A DISCOID
CC       ARRANGEMENT OF 5 NONCOVALENTLY BOUND SUBUNITS.
CC   -!- SIMILARITY: BELONGS TO THE PENTAXIN FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch)
CC   --------------------------------------------------------------------------
DR   EMBL; M14026; AAA28270.1; -.
DR   PIR; C25192; C25192.
DR   PIR; C25193; C25193.
DR   PDB; 1LIM; 15-OCT-95.
     InterPro; IPR001759; -.
     Pfam; PF00354; pentaxin; 1.
DR   PRINTS; PR00895; PENTAXIN.
DR   PROSITE; PS00289; PENTAXIN; 1.
KW   Pentaxin; Calcium; Glycoprotein; Signal; 3D-structure.
FT   SIGNAL        1     24
FT   CHAIN        25    242       C-REACTIVE PROTEIN 1.1.
FT   DOMAIN       25    242       PENTAXIN.
FT   DISULFID     62    125
FT   DISULFID    112    144
FT   DISULFID    207    241
FT   CARBOHYD    147    147       N-LINKED (GLCNAC. . . .).
FT   DOMAIN       52     67       BINDING OF PHOSPHORYLCHOLINE (POTENTIAL).
FT   DOMAIN       59     60       INTERACTION WITH PHOSPHATE GROUP OF
FT                                PHOSPHORYLCHOLINE (POTENTIAL)
FT   BINDING      63     63       QUATERNARY AMMONIUM GROUP OF
FT                                PHOSPHORYLCHOLINE (POTENTIAL).
FT   SIMILAR     139    153       STRONG WITH CA-BINDING EF-HAND SEQUENCE.
SQ   SEQUENCE   242 AA;  26773 MW;  A5[illegible] A8FDB92 CRC64;

  Query Match             66.0%;  Score [illegible];  DB [illegible];  Length 242;
  Best Local Similarity   100.0%;  Pred. No. 9;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    [illegible];  Gaps    [illegible];

Qy        3 HGPIC 7
            |||||
Db        5 HGPIC 9

Search completed: August 22, 2001, 14:30:[illegible]
Job time: 411 sec
```

```
                    GenCore version 4.5
          Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on:          August 22, 2001, 14:25:27  Search time 44.49 Seconds
                                             (without alignments)
                                             7.087 Million cell updates/sec

Title:           US-08 659-254-6
Perfect score:   38
Sequence:        1 NQEGWK 6

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters:      219241

[illegible]num DB seq length: 0
[illegible]mum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       PIR_68:*
                1:  pir1:*
                2:  pir2:*
                3:  pir3:*
                4:  pir4:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                              SUMMARIES
                  %
Result          Query
  No.   Score   Match Length DB  ID                      Description
------------ ---------------------------------- -- - --- ----------- --    -
     1     38   100.0    284  2  A55455                  steroidogenic acut
     2     38   100.0    285  2  I38248                  steroidogenic acut
     3     38   100.0    362  2  JC5386                  steroidogenic acut
     4     35    92.1    358  2  S31407                  (2'-5')oligo(A) sy
     5     35    92.1    368  2  T50174                  probable saccharop
     6     35    92.1    783  2  JC5467                  cellulase (EC 3.2.
     7     35    92.1    824  3  JC7532                  cellulase (EC 3.2.
     8     33    86.8    285  2  JC4315                  steroidogenic acut
     9     33    86.8    433  2  E83875                  hypothetical prote
    10     33    86.8    620  2  S64304                  hypothetical prote
    11     33    86.8    622  2  T23804                  hypothetical prote
    12     33    86.8   3587  2  I40486                  surfactin syntheta
    13     33    86.8   3588  2  I40485                  surfactin syntheta
    14     32    84.2    271  2  F64694                  hypothetical prote
    15     32    84.2    278  2  H64029                  hypothetical prote
    16     32    84.2    315  2  I39479                  hypothetical prote
    17     32    84.2    333  1  KHHUI                   cathepsin L (EC 3.
    18     32    84.2    334  2  A58195                  cathepsin L (EC 3.
    19     32    84.2    476  2  T20982                  hypothetical prote
    20     32    84.2    487  2  PQ0259                  hypothetical prote
    21     32    84.2    500  2  T10543                  hypothetical prote
    22     32    84.2    616  1  A46328                  hemagglutinin-neur
    23     32    84.2    699  1  XEBYUC                  UDPglucose 4-epime
    24     32    84.2    714  2  S68603                  hypothetical prote
    25     32    84.2    716  2  D69855                  conserved hypothet
    26     32    84.2    847  1  A53800                  mixed lineage prot
    27     32    84.2    977  2  C64809                  ybgB protein - Esc
    28     32    84.2   1125  2  T19193                  hypothetical prote
    29     32    84.2   1271  2  D64217                  hypothetical prote
    30     32    84.2   1943  2  B64596                  toxin like outer m
    31     32    84.2   2115  2  S38480                  nonstructural prot
    32     32    84.2   2205  1  MNWVRN                  nonstructural poly
    33     31    81.6    359  2  I59403                  alpha 2,8-polysial
    34     31    81.6    410  1  S28615                  serine/threonine/t
    35     31    81.6    412  2  T02725                  probable serine/th
    36     31    81.6    539  2  T27559                  hypothetical prote
    37     31    81.6    727  2  T23585                  hypothetical prote
    38     30    78.9    119  2  T08245                  gas-vesicle operon
    39     30    78.9    120  2  S28130                  gas vesicle operon
    40     30    78.9    131  2  E59097                  hypothetical prote
    41     30    78.9    131  2  D82222                  hypothetical prote
    42     30    78.9    230  2  T45059                  hypothetical prote
    43     30    78.9    253  2  H69088                  hypothetical prote
    44     30    78.9    259  2  T26298                  hypothetical prote
    45     30    78.9    280  2  F71077                  hypothetical prote

                              ALIGNMENTS

RESULT    1
A55455
steroidogenic acute regulatory protein precursor, mitochondrial - mouse
C;Species: Mus musculus (house mouse)
C;Date: 10-Feb 1995 #sequence_revision 10-Feb 1995 #text_change 08-Oct-1999
C;Accession: A55455
R;Clark, B.J.; Wells, J.; King, S.R.; Stocco, D.M.
J. Biol. Chem. 269, 28314 28322, 1994
A;Title: The purification, cloning, and expression of a novel luteinizing hormone-ind
ry protein (StAR).
A;Reference number: A55455; MUID:95050610
A;Accession: A55455
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-284 <CLA>
A;Cross-references: GB:L36062; GB:L36061; GB:S73839; NID:g1236242; PIDN:AAB94783.1; P
C;Keywords: mitochondrion; phosphoprotein
F;1-47/Domain: transit peptide (mitochondrion) #status predicted <TNP>

  Query Match              100.0%;  Score 38;  DB 2;  Length 284;
  Best Local Similarity    100.0%;  Pred. No. 6.2;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 NQEGWK 6
            ||||||
Db       91 NQEGWK 96

RESULT    2
I38248
steroidogenic acute regulatory protein - human
C;Species: Homo sapiens (man)
C;Date: 29 May 1998 #sequence_revision 29 May 1998 #text_change 05 Nov-1999
C;Accession: I38248; I38896
R;Sugawara, T.; Lin, D.; Holt, J.A.; Martin, K.O.; Javitt, N.B.; Miller, W.L.; Straus
Biochemistry 34, 12506 12512, 1995
A;Title: Structure of the human steroidogenic acute regulatory protein (StAR) gene: S
A;Reference number: I38248; MUID:96038[illegible]
A;Accession: I38248
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-285 <RES>
A;Cross references: EMBL:U29105; NID:g[illegible]; PIDN:AA[illegible]; PID:g1 41598
R;Sugawara, T.; Holt, J.A.; Driscoll, D.; Strauss III, J.F.; Lin, D.; Miller, W.L.; P
Proc. Natl. Acad. Sci. U.S.A. 92, 4778 4782, 1995
A;Title: Human steroidogenic acute regulatory protein: functional activity in COS 1 c
A;Reference number: I38896; MUID:95281[illegible]
A;Accession: I38896
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-285 <RE2>
A;Cross-references: EMBL [illegible]; NID [illegible]; PIDN:AA[illegible]
C;Genetics:
A;Gene: STAR
A;Cross-references: GDB: AR; GDB:[illegible]; MIM:[illegible]
A;Map position: 8p11.2-8p11.2
A;Introns: 22/1; [illegible]

  Query Match           [illegible]
  Best Local Similarity [illegible]; Pred. No. [illegible]
  Matches    6; Conservative  [illegible]; Mismatches [illegible]

Qy          1 NQEGWK 6
                ||||
Db         92 NQEGWK 97


[illegible]T   3
[illegible]6
ste[illegible]oidogenic acute regulatory protein - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 04-Jun-1997 #sequence_revision [illegible] #text_change [illegible]
C;Accession: JC5380
R;Lee, H.; Ahn, R.[illegible]; Kwo[illegible] H.; [illegible]
Biochem. Biophys. Res. Commun. 230, [illegible] 1997
A;Title: Nucleotide sequence [illegible]
A;Reference number: [illegible]; MUID:[illegible]
A;Accession: JC5380
A;Molecule type: mRNA
A;Residues: 1-362 [illegible]
A;Cross-references: [illegible]
C;Comment: This protein is a key regulator for the [illegible] acute response to t

  Query Match           [illegible]; Score [illegible]; DB [illegible]; Length [illegible]
  Best Local Similarity [illegible]; Pred. No. [illegible]
  Matches    6; Conservative    0; Mismatches [illegible]; Indels [illegible]; Gaps [illegible]

Qy          1 NQEGWK 6
              ||||||
Db        169 NQEGWK 174


RESULT    4
S31407
[illegible]5')oligo(A) synthetase (E[illegible]) - rat
[illegible]ernate names: 2'-5'-oligoadenylate synthetase
[illegible]pecies: Rattus norvegicus (Norway rat)
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change [illegible]
C;Accession: S31407
R;Truve, E.; Kelve, M.; Aaspollu, A.; Kuusksalu, A.; [illegible]
submitted to the EMBL Data Library, [illegible]
A;Description: Principles and background for the [illegible]
A;Reference number: S31407
A;Accession: S31407
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-358 [illegible]
A;Cross-references: EMBL:Z[illegible]; NID:[illegible]; PIDN:[illegible]
C;Superfamily: oligo(A) synthetase
C;Keywords: nucleotidyltransferase

  Query Match           [illegible]
  Best Local Similarity [illegible]; Pred. No. [illegible]
  Matches    5; Conservative [illegible]; Mismatches [illegible]

Qy          1 NQEGWK 6
              |||| :
Db        314 NQEGWR 319


RESULT    5
```

```
[illegible]
probable [illegible] - fission yeast (Schizosaccharomyces p
C;Species: Schizosaccharomyces pombe
C;Date: [illegible] #sequence_revision [illegible] #text_change [illegible]
C;Accession: [illegible]
R;Zimmermann, W.; [illegible]; McDonald, [illegible]; [illegible] M.A.; Barrell, B.G.
submitted to the EMBL Data Library, November [illegible]
A;Reference number: [illegible]
A;Accession: [illegible]
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: [illegible]
A;Cross-references: EMBL:[illegible]; PIDN:[illegible]; GSPDB:GN[illegible]
A;Experimental source: strain [illegible]
C;Genetics:
A;Gene: [illegible]
A;Map position: [illegible]

  Query Match           [illegible]; Score [illegible]; DB [illegible]; Length [illegible]
  Best Local Similarity [illegible]; Pred. No. [illegible]
  Matches    [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible]

Qy          [illegible] NQEGWK [illegible]

Db          [illegible] NQEGWK [illegible]


RESULT    [illegible]
[illegible]
[illegible] (EC [illegible]) - Bacillus sp.
N;Alternate names: Endo-1,4-beta-[illegible]anase
C;Species: Bacillus sp.
C;Date: 07-Jul-[illegible] #sequence_revision 18-[illegible] #text_change 10-Dec-1999
C;Accession: [illegible]
R;Miyatake, M.; Imada, K.
Biosci. Biotechnol. Biochem. [illegible], 1998
A;Title: A gene encoding 1,4-beta-[illegible] from Bacillus sp. [illegible]
A;Reference number: [illegible]; MUID:[illegible]
A;Accession: [illegible]
A;Molecule type: DNA
A;Residues: [illegible]
A;Cross-references: DDBJ:[illegible]
A;Experimental source: strain [illegible]
A;Note: [illegible] the complete translation are s
C;Superfamily: [illegible] alkaline cellulase; Thermotoga xylanase A amino-terminal
C;Keywords: [illegible]
F;[illegible]/Domain: [illegible] xylanase A amino-terminal repeat homology <XA>

  Query Match           [illegible]; Score [illegible]; DB [illegible]; Length [illegible]
  Best Local Similarity [illegible]; Pred. No. [illegible]
  Matches    [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible]

Qy          [illegible] NQEGWK [illegible]

Db          [illegible] NQEGWK [illegible]


RESULT    [illegible]
[illegible]
[illegible] (EC [illegible]) - Bacillus sp. (strain KSM-S237)
N;Alternate names: endo-1,4-beta-glucanase
C;Species: Bacillus sp. (strain KSM-S237)
C;Date: 31-Mar-2001 #sequence_revision 31-Mar-2001 #text_change 31-Mar-2001
C;Accession: [illegible]
R;Hakamada, Y.; Hatada, [illegible]; K[illegible]; Yoshimatsu, [illegible]; Kawai, S.; Kobayashi, T.; Ito,
Biosci. Biotechnol. Biochem. [illegible]
A;Title: [illegible] thermostabil
A;Reference number: [illegible]; MUID:[illegible]
A;Accession: [illegible]
A;Molecule type: DNA
A;Residues: [illegible]
```

```
A;Cross-references: DDBJ:AB018420
A;Experimental source: strain KSM-S237
A;Accession: PC7107
A;Molecule type: protein
A;Residues: 31-50 <HA2>
C;Comment: This enzyme is a thermostable, alkaline enzyme used as effective detergent ad
C;Genetics:
A;Gene: Egl-237
C;Keywords: hydrolase; glycosidase


  Query Match             92.1%;  Score 35;  DB 4;  Length 824;
  Best Local Similarity   83.3%;  Pred. No. [illegible];
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 NQEGWK 6
           |:||||
Db     204 NEEGWK 209


RESULT   8
JC4315
[illegible]oidogenic acute regulatory protein - bovine
[illegible]ecies: Bos primigenius taurus (cattle)
[illegible]ate: 29 Nov 1995 #sequence_revision 29 [illegible] 1996 #text_change [illegible]
C;Accession: JC4315
R;Hartung, S.; Rust, W.; Balvers, M.; Ivell, R.
Biochem. Biophys. Res. Commun. 215, 64[illegible]-65[illegible], 1995
A;Title: Molecular cloning and in vivo expression of the bovine steroidogenic acute regu
A;Reference number: JC4315; MUID:96011827
A;Accession: JC4315
A;Molecule type: mRNA
A;Residues: 1-285 <HAR>
C;Comment: This protein is an acute controller of the rate-limiting transfer of cholester
C;Genetics:
A;Gene: StAR
F;226-264/Region: metalloproteinase-1 tissue inhibitor similarity


  Query Match             86.8%;  Score [illegible];  DB 2;  Length 285;
  Best Local Similarity   83.3%;  Pred. No. [illegible]4;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 NQEGWK 6
           :|||||
Db      92 DQEGWK 97


RESULT   9
E83875
[illegible]thetical protein BH1805 [imported] - Bacillus halodurans (strain C-125)
[illegible]ecies: Bacillus halodurans
[illegible]te: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 01-Dec-2000
C;Accession: E83875
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A;Reference number: A83650; MUID:2026[illegible]314
A;Accession: E83875
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-433 <STO>
A;Cross-references: GB:AP001513; GB:BA0[illegible]; NID:g10174[illegible]45; PIDN:BAB0[illegible]524.1; GSPDB:GN00
A;Experimental source: strain C-125
C;Genetics:
A;Gene: BH1805


  Query Match             86.8%;  Score [illegible];  DB 2;  Length 433;
  Best Local Similarity   100.0%;  Pred. No. 83;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 NQEGW 5
```

```
           |||||
Db     375 NQEGW 379


RESULT  10
S64304
hypothetical protein YGR013w - yeast (Saccharomyces cerevisiae)
N;Alternate names: hypothetical protein G4 14
C;Species: Saccharomyces cerevisiae
C;Date: 17 May-1996 #sequence_revision 17 May 1996 #text_change 29 Oct-1999
C;Accession: S64304
R;Rieger, M.; Mueller Auer, S.; Brueckner, M.; Schaefer, M.
submitted to the Protein Sequence Database, May 1996
A;Reference number: S64071
A;Accession: S64304
A;Molecule type: DNA
A;Residues: 1-620 <RIE>
A;Cross references: EMBL:Z72798; NID:g13[illegible]297[illegible]; PID:e24[illegible]69; PID:g1322976; GSPDB:GN000
A;Experimental source: strain S288C
C;Genetics:
A;Gene: MIPS:YGR013w
A;Map position: 7R


  Query Match             86.8%;  Score [illegible];  DB 2;  Length 620;
  Best Local Similarity   83.3%;  Pred. No. 1.2e+02;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 NQEGWK 6
           :|||||
Db      15 SQEGWK 20


RESULT  11
T23804
hypothetical protein M195.3 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T23804
R;Burton, J.
submitted to the EMBL Data Library, October 1995
A;Reference number: Z19801
A;Accession: T23804
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-622 <WIL>
A;Cross-references: EMBL:Z66498; PIDN:CAA91299.1; GSPDB:GN00020; CESP:M195.3
A;Experimental source: clone M195
C;Genetics:
A;Gene: CESP:M195.3
A;Map position: 2
A;Introns: [illegible]


  Query Match             86.8%;  Score [illegible];  DB 2;  Length 622;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 NQEGW 5
           |||||
Db     127 NQEGW 131


RESULT  12
I40486
surfactin synthetase component II - Bacillus subtilis
N;Alternate names: surfactin synthetase SrfAB; surfactin synthetase/competence protei
N;Contains: acid--amino acid ligase (EC 6.3.2.-)
C;Species: Bacillus subtilis
C;Date: 12 Aug 1996 #sequence_revision 12 Aug 1996 #text_change 03-Nov-2000
C;Accession: I40486; S60866; C69718; S40363; S35518; S25658; S34986
R;Cosmina, P.; Rodriguez, F.; de Ferra, F.; Grandi, G.; Perego, M.; Venema, G.; van S
Mol. Microbiol. 8, 821-831, 1993
```

```
A;Title: Sequence and analysis of the genetic locus responsible for surfactin synthes
A;Reference number: [illegible]104[illegible]; MUID:[illegible]
A;Accession: I40486
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-358[illegible]
A;Cross-references: EMBL:X7[illegible]; NID:[illegible]; PIDN:[illegible]
A;Experimental source: strain W168 derivative of [illegible]
R;Hamoen, L.W.; Eshuis, H.; Jongbloed, J.; Venema, [illegible]
Mol. Microbiol. 15, [illegible] 1995
A;Title: A small gene, designated [illegible], located within the [illegible]
A;Reference number: [illegible]; MUID:[illegible]
A;Accession: S6086[illegible]
A;Molecule type: DNA
A;Residues: 977-1[illegible] HAM
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; [illegible]
C[illegible]ron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.;
A[illegible]rlich, S.D.; Emmerson, P.T.; Entian, K.D.; [illegible]; Ferrari, E.
N[illegible]e 390, 249-256, 1997
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler
[illegible]rech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; [illegible]; Masuda, S.; Mauel
y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
 Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; [illegible] M.; Sadaie, Y.; Sato, T.; Scanlan,
A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; [illegible]; Sekowska, A.; Ser
ikeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; [illegible]; Tosato, V.; Uchiyama,
I.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis
A;Reference number: A[illegible]58; MUID:98044[illegible]
A;Accession: [illegible]6972[illegible]
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-3587 <KUN>
A;Cross-references: [illegible]Z99[illegible]; GB:AL[illegible]; NID:[illegible]; PIDN:[illegible]; PID:[illegible]
A;Experimental source: strain 168
R;Fabret, C.; Quentin, Y.; Guiseppi, A.; Busuttil, [illegible]
submitted to the EMBL Data Library, Ma[illegible]
A;Reference number: S4696[illegible]
A;Accession: S4696[illegible]
A;Molecule type: DNA
A;Residues: 1-32,'[illegible]',[illegible] [illegible]
1756-1914,'PK',1917-2138,'[illegible]',2142-[illegible] [illegible]
A;Cross-references: EMBL:X[illegible]; NID:[illegible]; PIDN:[illegible]
[illegible]a, S.; Fujishima, Y.; Corbell, N.; [illegible]; Nakano, M.M.; [illegible]; Yamane, K.
[illegible]ic Acids Res. 21, [illegible] 1993
A;Title: Nucleotide sequence of [illegible] portion of srfA that [illegible] required [illegible]
A;Reference number: S[illegible]; MUID:[illegible]
A;Accession: S3551[illegible]
A;Status: significant sequence differences
A;Molecule type: DNA
A;Cross-references: EMBL:[illegible]; NID:[illegible]
A;Experimental source: strain 168 trpC2
R;Borchert, S.; Patil, S.S.; Marahiel, M.A.
FEMS Microbiol. Lett. 92, [illegible]-180, 199[illegible]
A;Title: Identification of putative multifunctional [illegible]
A;Reference number: S[illegible]658
A;Accession: S2565[illegible]
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 168,'[illegible]
A;Cross-references: EMBL:X[illegible]; NID:[illegible]; PIDN:[illegible]
A;Experimental source: strain A[illegible]
C;Comment: This protein contains several [illegible]
the amino-terminal region of this protein appear to be required for the [illegible] of
C;Genetics:
A;Gene: srfAB; srfA2
C;Superfamily: surfactin synthetase [illegible]
C;Keywords: antibiotic biosynthesis; carrier protein; [illegible]
F;[illegible]1-951/Domain: acetate-[illegible]A ligase homology <A[illegible]
F;968-1035/Domain: acyl carrier protein homology <A[illegible]
F;1036-1481/Domain: repeat <RPT[illegible]
F;[illegible]4[illegible]/Domain: acetate-[illegible]A ligase homology <A[illegible]
F;[illegible]/Domain: acyl carrier protein homology <ACP[illegible]
F;[illegible]/Domain: repeat <RPT[illegible]
F;[illegible]/Domain: acetate-CoA ligase homology <ACL[illegible]
F;[illegible]/Domain: acyl carrier protein homology <A[illegible]
F;[illegible]/Binding site: phosphopantetheine (Ser) (covalent) #status predicted

  [illegible] Match [illegible]; Score [illegible]; [illegible] Length [illegible];
  Best [illegible]; [illegible]; [illegible]
  Matches [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible]

Qy          1 N[illegible]WK[illegible]

Db       [illegible]984 N[illegible]WK[illegible]


RESULT [illegible]
I4[illegible]48[illegible]
surfactin synthetase component [illegible] - Bacillus subtilis
N;Alternate names: competence protein srfAA; surfactin production protein srfAA; surf
C;Species: Bacillus subtilis
C;Date: [illegible]; #sequence_revision [illegible]; #text_change [illegible] Nov 2000
C;Accession: I4[illegible]48[illegible]; B[illegible]; S[illegible]; A[illegible]; [illegible]; A4[illegible]; S4[illegible]
R;Cosmina, P.; Rodriguez, F.; de Ferra, F.; Grandi, G.; Perego, M.; Venema, G.; van S
Mol. Microbiol. [illegible], [illegible] 199[illegible]
A;Title: Sequence and analysis of the genetic locus responsible for surfactin synthes
A;Reference number: I4[illegible]; MUID:[illegible]
A;Accession: I4048[illegible]
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: [illegible] <RES>
A;Cross-references: EMBL:X[illegible]; NID:[illegible]; PIDN:AA[illegible]; PID:[illegible]
A;Experimental source: strain W168 derivative of JH642
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Ber
C.; Brignell, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.;
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari,
Nature 390, 249-256, 1997
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gal
[illegible]rech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M
Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardino
A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mau
y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portete
 Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanl
A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Se
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiya
I.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida
A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtil
A;Reference number: A[illegible]; MUID:98044[illegible]
A;Accession: [illegible]
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: [illegible] <KUN>
A;Cross-references: [illegible]; GB:A[illegible]; NID:[illegible]4957; PIDN:AB12[illegible]; PID:[illegible]
A;Experimental source: strain 168
R;Fuma, S.; Fujishima, Y.; Corbell, N.; D'Souza, C.; Nakano, M.M.; Zuber, P.; Yamane
Nucleic Acids Res. 21, [illegible] 1993
A;Title: Nucleotide sequence of [illegible] portion of srfA that contains the [illegible] required
A;Reference number: [illegible]; MUID:[illegible]
A;Accession: [illegible]
A;Status: nucleic acid sequence not shown; significant sequence differences
A;Molecule type: DNA
A;Cross-references: EMBL:[illegible]; NID:[illegible]16345; PID:[illegible]16346
A;Experimental source: strain [illegible]
A;Note: protein sequence not complete, the nucleotide sequence was identified [illegible]
R;Nakano, M.M.; Magnuson, R.; Myers, A.; Curry, J.; Grossman, A.D.; Zuber, P.
J. Bacteriol. [illegible]
A;Title: srfA is an operon required for surfactin production, competence development
A;Reference number: A[illegible]; MUID:[illegible]
A;Accession: A[illegible]
A;Status: [illegible] not shown; not compared with conceptual translation
A;Molecule type: DNA
```

```
A;Residues: 1-46,'EV',49-145,'I',147 15[illegible],'[illegible]',[illegible]52,'AN',155 280,'I',282 [illegible]7,'[illegible]',[illegible]0 [illegible]84
A;Cross-references: GB:M59939; NID:g14366
R;Nakano, M.M.; Xia, L.; Zuber, P.
J. Bacteriol. 173, 5487-5493, 1991
A;Title: Transcription initiation region of the srfA operon, which is controlled by the
A;Reference number: A43705; MUID:91358326
A;Contents: annotation
R;Fabret, C.; Quentin, Y.; Guiseppi, A.; Busuttil, J.; Haiech, J.; Denizot, F.
submitted to the EMBL Data Library, March 199[illegible]
A;Reference number: S46967
A;Accession: S46967
A;Molecule type: DNA
A;Residues: 3249-3271,'A',3273 3316,'R',33[illegible] [illegible]451,'Y',345[illegible] 34[illegible],'[illegible]',4[illegible] [illegible]4[illegible],'DAH',34
A;Cross-references: EMBL:X72672; NID:g516[illegible]; PIDN:CAA51222.1; PID:g516[illegible]59
A;Experimental source: strain 168 trpC2
C;Comment: This enzyme is one of several in the multienzyme complex that synthesizes the
nce.
C;Genetics:
A;Gene: srfAA; srfA1
A;Note: srfAA is the first gene of the srfA operon
C;Superfamily: surfactin synthetase; acetate--CoA ligase homology; acyl carrier protein
C;Keywords: antibiotic biosynthesis; carrier protein; phosphopantetheine; phosphoprotein
 7-956/Domain: acetate--CoA ligase homology <ACL1>
 4-1042/Domain: acyl carrier protein homology <ACP1>
F 043-1488/Domain: repeat <RPT1>
F;1549-1993/Domain: acetate--CoA ligase homology <ACL2>
F;2011-2079/Domain: acyl carrier protein homology <ACP2>
F;2080-2527/Domain: repeat <RPT2>
F;2589-3025/Domain: acetate--CoA ligase homology <ACL3>
F;3042-3109/Domain: acyl carrier protein homology <ACP3>
F;1006,2043,3074/Binding site: phosphopantetheine (Ser) (covalent) #status predicted


  Query Match             86.8%;  Score 33;  DB 2;  Length 3588;
  Best Local Similarity   83.3%;  Pred. No. [illegible]5e-02;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 NQEGWK 6
              || |||
Db       3085 NQHGWK 3090


RESULT  14
E64694
hypothetical protein HP1397 - Helicobacter pylori (strain 26695)
C;Species: Helicobacter pylori
C;Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 08-Oct-1999
C;Accession: E64694
R;Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.
 Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenne
s   J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L.
  re 388, 539-547, 1997
  thors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.
A;Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
A;Reference number: A64520; MUID:97394467
A;Accession: E64694
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-271 <TOM>
A;Cross-references: GB:AE000639; GB:AE000511; NID:g2314560; PIDN:AAD0844[illegible].1; PID:g231457

  Query Match             84.2%;  Score 32;  DB 2;  Length 271;
  Best Local Similarity   100.0%;  Pred. No. 78;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 QEGWK 6
              |||||
Db        111 QEGWK 115


RESULT  15
H64029
```

```
hypothetical protein HI1427 - Haemophilus influenzae (strain Rd KW20)
C;Species: Haemophilus influenzae
C;Date: 18 Aug 1995 #sequence_revision 18 Aug 1995 #text_change 04-Mar-2000
C;Accession: H64029
R;Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage
; Cougayne, J.D.; Scott, J.; Shirley, R.; [illegible], L.I.; Glodek, A.; Kelley, J.M.; Weidman
, D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.
Science 269, 496-512, 1995
A;Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Vente
A;Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
A;Reference number: A64000; MUID:95350630
A;Accession: H64029
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-278 <TIGR>
A;Cross-references: GB:U32821; GB:L42023; NID:g1574240; PIDN:AAC23071.1; PID:g1574261
C;Superfamily: Haemophilus influenzae hypothetical protein HI1427


  Query Match             84.2%;  Score 32;  DB 2;  Length 278;
  Best Local Similarity   83.3%;  Pred. No. 80;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 NQEGWK 6
              | ||||
Db        107 NAEGWK 112



Search completed: August 22, 2001, 4:2[illegible]:[illegible]3
Job time: 82 sec
```

```
                    GenCore version 4.5
          Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on:          August 22, 2001, 14:40:56 ; Search time 35.8 Seconds
                                             (without alignments)
                                             5.741 Million cell updates/sec

Title:           US-08-659-254-6
Perfect score:   38
Sequence:        1 NQEGWK 6

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters:      93435

 mum DB seq length: 0
 mum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       SwissProt_39:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                              SUMMARIES

              %
Result      Query
  No   Score Match Length DB  ID                          Description
-------------------------------------------------------------------------
   1     38  100.0    284  1  STAR_MESAU                  P70114 mesocricetu
   2     38  100.0    284  1  STAR_MOUSE                  P51557 mus musculu
   3     38  100.0    284  1  STAR_RAT                    P97826 rattus norv
   4     38  100.0    285  1  STAR_HORSE                  O46689 equus cabal
   5     38  100.0    285  1  STAR_HUMAN                  P49675 homo sapien
   6     38  100.0    285  1  STAR_PIG                    Q28996 sus scrofa
   7     38  100.0    475  1  TBE_HUMAN                   Q9ujt0 homo sapien
   8     35   92.1    358  1  25A1_RAT                    Q05961 rattus norv
   9     35   92.1    358  1  OAS1_RAT                    Q05961 rattus norv
  10     35   92.1    368  1  LYS1_SCHPO                  Q09694 schizosacch
  11     33   86.8    285  1  STAR_BOVIN                  Q28918 bos taurus
  12     33   86.8    514  1  GUXC_FUSOX                  P46239 fusarium ox
  13     33   86.8    620  1  YG18_YEAST                  P53207 saccharomyc
  14     33   86.8    622  1  AMT3_CAEEL                  Q21565 caenorhabdi
  15     33   86.8   3587  1  SRF1_BACSU                  P27206 bacillus su
  16     33   86.8   3587  1  SRF2_BACSU                  Q04747 bacillus su
  17     32   84.2    278  1  YE27_HAEIN                  P44196 haemophilus
  18     32   84.2    333  1  CATL_HUMAN                  P07711 homo sapien
  19     32   84.2    334  1  CATL_PIG                    Q28944 sus scrofa
  20     32   84.2    334  1  CATM_HUMAN                  O60911 homo sapien
  21     32   84.2    616  1  HEMA_NDVI                   P12555 newcastle d
  22     32   84.2    699  1  CALX_YEAST                  P27825 saccharomyc
  23     32   84.2    977  1  YBCG_ECOLI                  P54746 escherichia
  24     32   84.2   1271  1  Y338_MYCPN                  P47580 mycoplasma
  25     32   84.2   2205  1  POLN_RUBVT                  P13889 rubella vir
  26     31   81.6    359  1  PST_HUMAN                   Q92187 homo sapien
  27     31   81.6    410  1  APKA_ARATH                  Q06548 arabidopsis
  28     31   81.6    412  1  APKB_ARATH                  P46573 arabidopsis
  29     30   78.9    119  1  GVO1_HALN1                  O51968 halobacteri
  30     30   78.9    120  1  GVO1_HALN2                  Q48313 halobacteri
  31     30   78.9    253  1  HEM4_METTH                  O26268 methanobact
  32     30   78.9    269  1  GEM2_RAT                    Q9qzp1 rattus norv
  33     30   78.9    280  1  GEM2_HUMAN                  O14893 homo sapien
  34     30   78.9    421  1  YBIU_ECOLI                  P75791 escherichia
  35     30   78.9    452  1  AAMP_HUMAN                  Q13685 homo sapien
  36     30   78.9    461  1  FCMD_HUMAN                  O75072 homo sapien
  37     30   78.9    498  1  YAOI_SCHPO                  Q10067 schizosacch
  38     30   78.9    505  1  TCMO_ARATH                  P92994 arabidopsis
  39     30   78.9    631  1  NTP1_VACCA                  O57214 vaccinia vi
  40     30   78.9    631  1  NTP1_VACCC                  P20637 vaccinia vi
  41     30   78.9    631  1  NTP1_VACCV                  P05807 vaccinia vi
  42     30   78.9    631  1  NTP1_VARV                   P33066 variola vir
  43     30   78.9    800  1  GUN_BACSP                   P06564 bacillus sp
  44     30   78.9    890  1  NIA2_PHAVU                  P39866 phaseolus v
  45     30   78.9    900  1  NIA_LOTJA                   P39869 lotus japon

                              ALIGNMENTS

RESULT   1
STAR_MESAU
ID   STAR_MESAU     STANDARD;      PRT;   284 AA.
AC   P70114;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN PRECURSOR.
GN   STAR.
OS   Mesocricetus auratus (Golden hamster).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC   Mesocricetus.
OX   NCBI_TaxID=10036;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97124769; PubMed 8969905;
RA   Fleury A., Cloutier M., Ducharme L., Lefebvre A., Lehoux J.,
RA   Lehoux J.-G.;
RT   "Adrenocorticotropin regulates the level of the steroidogenic acute
RT   regulatory (StAR) protein mRNA in hamster adrenals.";
RL   Endocr. Res. 22:515-520(1996).
CC   -!- FUNCTION: PLAYS A KEY ROLE IN STEROID HORMONE SYNTHESIS BY
CC       ENHANCING THE METABOLISM OF CHOLESTEROL INTO PREGNENOLONE.
CC       TRANSPORTER THAT BINDS TO AND TRANSPORT CHOLESTEROL THROUGH THE
CC       INTERMEMBRANE SPACE OF THE MITOCHONDRION (PROBABLE).
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL.
CC   -!- SIMILARITY: BELONGS TO THE STAR FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are  no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; U66490; AAB06763.1; -.
DR   InterPro; IPR000799; -.
DR   InterPro; IPR002913; -.
DR   Pfam; PF01852; START; 1.
DR   PRINTS; PR00978; STARPROTEIN.
KW   Lipid-binding; Lipid transport; Transport; Steroidogenesis;
KW   Mitochondrion; Transit peptide.
FT   TRANSIT       1      ?       MITOCHONDRION (POTENTIAL).
FT   CHAIN         ?    284       STEROIDOGENIC ACUTE REGULATORY PROTEIN.
SQ   SEQUENCE   284 AA;  31497 MW;  [illegible] CRC64;

  Query Match             100.0%;  Score 38;  DB 1;  Length 284;
  Best Local Similarity   100.0%;  Pred. No. 2.5;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 NQEGWK 6
            ||||||
Db       91 NQEGWK 96
```

```
RESULT 2
STAR MOUSE
ID   STAR MOUSE     STANDARD;     PRT;   284 AA.
AC   P51557;
DT   01 OCT 1996 (Rel. 34, Created)
DT   01 OCT 1996 (Rel. 34, Last sequence update)
DT   30 MAY 2000 (Rel. 39, Last annotation update)
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN (PRECURSOR) [illegible] HORMONE
DE   INDUCED PROTEIN)
GN   STAR.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; [illegible]
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID [illegible]
RN   [1]
RP   SEQUENCE FROM N.A. [illegible] PARTIAL [illegible]
RX   MEDLINE 9505[illegible]; PubMed [illegible]
RA   Clark B.J., Wells J., King S.R., [illegible] M.;
RT   "The purification, cloning, and expression of [illegible]
RT   hormone-induced mitochondrial protein in MA[illegible]
RT   cells. Characterization of the steroidogenic acute regulatory protein
RT   (STAR).";
RL   J. Biol. Chem. [illegible] (1994).
CC   -!- FUNCTION: PLAYS A KEY ROLE IN STEROID HORMONE SYNTHESIS BY
CC       ENHANCING THE METABOLISM OF CHOLESTEROL INTO PREGNENOLONE.
CC       TRANSPORTER THAT BINDS TO AND TRANSFERS [illegible]
CC       INTERMEMBRANE SPACE OF THE MITOCHONDRION (PROBABLE).
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL.
CC   -!- SIMILARITY: BELONGS TO THE STAR FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; L36062; AAB[illegible]
DR   MGD; MGI:102760; Star.
DR   InterPro; IPR00[illegible];
DR   InterPro; IPR00[illegible];
DR   Pfam; PF01852; START; 1.
DR   PRINTS; PR00978; STARPROTEIN.
KW   Lipid-binding; Lipid transport; Transport; Steroidogenesis;
KW   Mitochondrion; Transit peptide.
FT   TRANSIT       1     [illegible]       MITOCHONDRION (POTENTIAL).
FT   CHAIN        [illegible]   284       STEROIDOGENIC ACUTE REGULATORY PROTEIN.
FT   DOMAIN       [illegible]   [illegible]       POLY-ALA.
SQ   SEQUENCE   284 AA;  [illegible] MW;  [illegible] CRC64;

  Query Match             [illegible]%;  Score [illegible];  DB [illegible];  Length 284;
  Best Local Similarity   [illegible]%;  Pred. No. [illegible];
  Matches  [illegible];  Conservative  [illegible];  Mismatches  [illegible];  Indels  [illegible];  Gaps  [illegible];

Qy          1 NQEGWK 6
              ||||||
Db         91 NQEGWK 96


RESULT 3
STAR RAT
ID   STAR RAT     STANDARD;     PRT;   284 AA.
AC   P97826; P97628;
DT   01 NOV 1997 (Rel. 35, Created)
DT   01 NOV 1997 (Rel. 35, Last sequence update)
DT   30 MAY 2000 (Rel. 39, Last annotation update)
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN (PRECURSOR) [illegible]
GN   STAR.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID [illegible]
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=WISTAR; TISSUE=Ovary;
RX   MEDLINE [illegible]; PubMed [illegible];
RA   Mizutani T., Sonoda Y., Minegishi T., Wakabayashi K., Miyamoto K.;
RT   "Molecular cloning, characterization and cellular distribution of rat
RT   steroidogenic acute regulatory protein (STAR) in the ovary.";
RL   Life Sci. [illegible]:[illegible]-[illegible](1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY; TISSUE=Testis;
RX   MEDLINE [illegible]; PubMed [illegible];
RA   Lee H.K., Ahn R.S., Kwon H.B., Soh J.;
RT   "Nucleotide sequence of the rat steroidogenic acute regulatory protein
RT   complementary DNA.";
RL   [illegible]
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY; TISSUE=Adrenal gland;
RX   MEDLINE [illegible]; PubMed [illegible];
RA   [illegible] [illegible], [illegible] Artemenko I., Bhattacharyya K.K.,
RA   Jefcoate C.R.;
RT   "Characterization of the rat StAR gene that encodes the predominant
RT   3.5-kilobase pair mRNA. ACTH stimulation of adrenal steroids in vivo
RT   precedes elevation of StAR mRNA and protein.";
RL   J. Biol. Chem. 27[illegible]:[illegible]-[illegible](1998).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY; TISSUE=Adrenal gland;
RX   MEDLINE [illegible]; PubMed [illegible];
RA   Furukawa A., Miyatake A., [illegible] [illegible], [illegible] Y.;
RT   "Steroidogenic acute regulatory protein (STAR) transcripts
RT   [illegible] expressed in the adult rat central nervous system:
RT   colocalization of STAR, cytochrome P-450SCC (CYPXIA1), and 3beta-
RT   hydroxysteroid dehydrogenase in the rat brain.";
RL   J. Neurochem. [illegible]:[illegible]-2238(1998).
CC   -!- FUNCTION: PLAYS A KEY ROLE IN STEROID HORMONE SYNTHESIS BY
CC       ENHANCING THE METABOLISM OF CHOLESTEROL INTO PREGNENOLONE.
CC       TRANSPORTER THAT BINDS TO AND TRANSFERS CHOLESTEROL THROUGH THE
CC       INTERMEMBRANE SPACE OF THE MITOCHONDRION (PROBABLE).
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL.
CC   -!- SIMILARITY: BELONGS TO THE STAR FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AB[illegible]; BAA[illegible]
DR   EMBL; [illegible]; AAB[illegible]
DR   EMBL; AF[illegible]; AA[illegible]
DR   EMBL; AB[illegible]; BAA[illegible]
DR   InterPro; IPR00[illegible];
DR   InterPro; IPR00[illegible];
DR   Pfam; PF01852; START; 1.
DR   PRINTS; PR00978; STARPROTEIN.
KW   Lipid-binding; Lipid transport; Transport; Steroidogenesis;
KW   Mitochondrion; Transit peptide.
FT   TRANSIT       1     [illegible]       MITOCHONDRION (POTENTIAL).
FT   CHAIN        [illegible]   284       STEROIDOGENIC ACUTE REGULATORY PROTEIN.
FT   CONFLICT      [illegible]   [illegible]       MLLATFK -> MIWPFLAVLFGLYVPERRPPTGMIGVGLPIN
FT                                   [illegible]VYFRLLQEFFFPMLESPNEIGSSGLVVVNRYLAFQEFAKR
FT                                   SYPPSEVLAYFLPC (IN REF. 2).
SQ   SEQUENCE   284 AA;  [illegible] MW;  [illegible] CRC64;


  Query Match             [illegible]%;  Score [illegible];  DB [illegible];  Length 284;
  Best Local Similarity   100.0%;  Pred. No. 2.5;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 NQEGWK 6
              ||||||
Db         91 NQEGWK 96


RESULT    4
STAR_HORSE
ID   STAR_HORSE     STANDARD;      PRT;   285 AA.
AC   O46689;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN PRECURSOR.
GN   STAR.
OS   Equus caballus (Horse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX   NCBI_TaxID=9796;
RN   [1]
[illegible]   SEQUENCE FROM N.A.
[illegible]   TISSUE=Follicle;
RX   MEDLINE=99124482; PubMed=9927292;
RA   Kerban A., Boerboom D., Sirois J.;
RT   "Human chorionic gonadotropin induces an inverse regulation of
RT   steroidogenic acute regulatory protein messenger ribonucleic acid in
RT   theca interna and granulosa cells of equine preovulatory follicles.";
RL   Endocrinology 140:667-674(1999).
CC   -!- FUNCTION: PLAYS A KEY ROLE IN STEROID HORMONE SYNTHESIS BY
CC       ENHANCING THE METABOLISM OF CHOLESTEROL INTO PREGNENOLONE.
CC       TRANSPORTER THAT BINDS TO AND TRANSPORT CHOLESTEROL THROUGH THE
CC       INTERMEMBRANE SPACE OF THE MITOCHONDRION (PROBABLE).
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIA..
CC   -!- SIMILARITY: BELONGS TO THE STAR FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as  long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AF031697; AAC04704.1; -.
DR   EMBL; AF031696; AAC04703.1; -.
DR   InterPro; IPR000799; -.
DR   InterPro; IPR002913; -.
DR   Pfam; PF01852; START; 1.
[illegible]   PRINTS; PR00978; STARPROTEIN.
[illegible]   Lipid-binding; Lipid transport; Transport; Steroidogenesis;
KW   Mitochondrion; Transit peptide.
FT   TRANSIT       1      ?       MITOCHONDRION (POTENTIAL).
FT   CHAIN         ?    285       STEROIDOGENIC ACUTE REGULATORY PROTEIN.
SQ   SEQUENCE   285 AA;  31852 MW;  B0360EF93299727D CRC64;


  Query Match             100.0%;  Score 38;  DB 1;  Length 285;
  Best Local Similarity   100.0%;  Pred. No. 2.5;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 NQEGWK 6
              ||||||
Db         92 NQEGWK 97


RESULT    5
STAR_HUMAN
ID   STAR_HUMAN     STANDARD;      PRT;   285 AA.
AC   P49675; Q16396;
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN PRECURSOR
GN   STAR.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Adrenal cortex;
RX   MEDLINE=95281540; PubMed=7761400;
RA   Sugawara T., Holt J.A., Driscoll D., Strauss J.F. III, Lin D.,
RA   Miller W.L., Patterson D., Clancy K.P., Hart I.M., Clark B.J.,
RA   Stocco D.M.;
RT   "Human steroidogenic acute regulatory protein: functional activity in
RT   COS-1 cells, tissue-specific expression, and mapping of the
RT   structural gene to 8p11.2 and a pseudogene to chromosome 13.";
RL   Proc. Natl. Acad. Sci. U.S.A. 92:4778-4782(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Placenta;
RX   MEDLINE=96038298; PubMed=7547998;
RA   Sugawara T., Lin D., Holt J.A., Martin K.O., Javitt N.B., Miller W.L.,
RA   Strauss J.F. III;
RT   "Structure of the human steroidogenic acute regulatory protein (STAR)
RT   gene: StAR stimulates mitochondrial cholesterol 27-hydroxylase
RT   activity.";
RL   Biochemistry 34:12506-12512(1995).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96008588; PubMed=7548191;
RA   Gradi A., Tang Wai R., McBride H.M., Chu L.L., Shore G.C.,
RA   Pelletier J.;
RT   "The human steroidogenic acute regulatory (StAR) gene is expressed in
RT   the urogenital system and encodes a mitochondrial polypeptide.";
RL   Biochim. Biophys. Acta 1258:228-233(1995).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RA   Yu W., Sarginson J., Gibbs R.A.;
RL   Submitted (NOV-1997) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   VARIANT CLAH LEU-182.
RX   MEDLINE=97096837; PubMed=8948562;
RA   Bose H.S., Sugawara T., Strauss J.F. III, Miller W.L.;
RT   "The pathophysiology and genetics of congenital lipoid adrenal
RT   hyperplasia.";
RL   New Engl. J. Med. 335:1870-1878(1996).
RN   [6]
RP   VARIANT ALA 203.
RX   MEDLINE=98112495; PubMed=9452116;
RA   Katsumata N., Tanae A., Shinagawa T., Nagashima M., Yokawa A.,
RA   Shimizu M., Yasunaga T., Tanaka T., Hibi I.;
RT   "A novel frameshift mutation 840delA and a novel polymorphism D203A in
RT   the steroidogenic acute regulatory protein gene in a Japanese patient
RT   with congenital lipoid adrenal hyperplasia.";
RL   Hum. Mutat. Suppl. 1:S304-S307(1998).
CC   -!- FUNCTION: PLAYS A KEY ROLE IN STEROID HORMONE SYNTHESIS BY
CC       ENHANCING THE METABOLISM OF CHOLESTEROL INTO PREGNENOLONE.
CC       TRANSPORTER THAT BINDS TO AND TRANSPORT CHOLESTEROL THROUGH THE
CC       INTERMEMBRANE SPACE OF THE MITOCHONDRION (PROBABLE).
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIA..
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN GONADS, ADRENAL CORTEX AND
CC       KIDNEY.
CC   -!- DISEASE: DEFECTS IN STAR CAUSE CONGENITAL LIPOID ADRENAL
CC       HYPERPLASIA (CLAH), THE MOST SEVERE FORM OF ADRENAL HYPERPLASIA.
CC   -!- SIMILARITY: BELONGS TO THE STAR FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as  long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.[illegible]
CC   or send an email to license@isb-sib.ch).
CC   -
DR   EMBL; U17280; AA[illegible]
DR   EMBL; U29105; AA[illegible]
DR   EMBL; U29099; AA[illegible]; [illegible]
DR   EMBL; U29100; AA[illegible]; [illegible]
DR   EMBL; U29101; AA[illegible]; [illegible]
DR   EMBL; U2910[illegible]; AA[illegible]; [illegible]
DR   EMBL; U29103; AA[illegible]; [illegible]
DR   EMBL; U29104; AA[illegible]; [illegible]
DR   EMBL; S79660; AAB35[illegible]
DR   EMBL; AF03[illegible]; AAB[illegible]
DR   MIM; 600617; -.
DR   MIM; 201710; -.
DR   InterPro; IPR00[illegible]
DR   InterPro; IPR0[illegible]
     Pfam; PF01852; START; 1.
     PRINTS; PR00978; STARPROTEIN.
KW   Lipid binding; Lipid transport; Transport; [illegible]
KW   Mitochondrion; Transit peptide; [illegible]
FT   TRANSIT    [illegible]        MITOCHONDRION (POTENTIAL).
FT   CHAIN      [illegible]        STEROIDOGENIC ACUTE [illegible]
FT   VARIANT    [illegible]        [illegible] (IN CAH[illegible]).
FT                                 [illegible] VAR[illegible]
FT   VARIANT    [illegible]        [illegible] A[illegible]
FT                                 [illegible] VAR[illegible]
SQ   SEQUENCE   285 AA;  [illegible] MW;  [illegible]

  Query Match            [illegible]
  Best Local Similarity  [illegible]
  Matches    [illegible]; Conservative [illegible]; Mismatches [illegible]

Qy          1 NQEGWK 6
              ||||||
Db         [illegible]2 NQEGWK [illegible]


RESULT 6
STAR_PIG
ID   STAR_PIG       STANDARD;      PRT;   [illegible] AA.
AC   Q28996; [illegible]
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
     30-MAY-2000 (Rel. 39, Last annotation update)
     STEROIDOGENIC ACUTE REGULATORY PROTEIN [illegible]
     STAR.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; [illegible]
OC   Mammalia; Eutheria; Cetartiodactyla; [illegible]
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Ovary;
RX   MEDLINE=97200784; PubMed=904861[illegible];
RA   Pilon N., Daneau I., [illegible]
RA   Silversides D.W.;
RT   "Porcine and bovine steroidogenic acute regulatory protein (StAR)
RT   gene expression during gestation.";
RL   Endocrinology [illegible]
RN   [2]
RP   SEQUENCE OF [illegible] FROM N.A.
RC   TISSUE=Corpus luteum;
RX   MEDLINE=971[illegible]; PubMed=[illegible];
RA   Balasubramanian K., [illegible] H.A., [illegible]
RA   Veldhuis J.D.;
RT   "Regulation of porcine granulosa cell [illegible]
RT   protein (StAR) [illegible]
RT   follicle-stimulating hormone or protein kinase A [illegible]."
RL   Endocrinology [illegible]
CC   -!- FUNCTION: PLAYS A KEY ROLE IN STEROID HORMONE SYNTHESIS [illegible]
CC       ENHANCING THE METABOLISM [illegible]
```

```
CC       TRANSFER [illegible] AND TRANSFER CHOLESTEROL THROUGH THE
CC       INTERMEMBRANE [illegible] (PROBABLE).
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN [illegible]
CC       TISSUES, [illegible] AND [illegible] ADRENAL
CC       [illegible] PREGNANCY, [illegible]
CC       DEVELOPING FETAL TISSUES, CORPUS LUTEUM AND PREGNANCY, [illegible] IN
CC       THE [illegible]
CC   -!- SIMILARITY: BELONGS TO THE STAR FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between [illegible] and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use [illegible] was
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; [illegible]; AAB[illegible]
DR   EMBL; [illegible]; AAB[illegible]
DR   InterPro; IPR00[illegible]
DR   InterPro; IPR0[illegible]
DR   Pfam; PF01852; START; 1.
DR   PRINTS; PR00978; STARPROTEIN.
KW   Lipid binding; Lipid transport; Transport; Steroidogenesis;
KW   Mitochondrion; Transit peptide.
FT   TRANSIT    [illegible]        MITOCHONDRION (POTENTIAL).
FT   CHAIN      [illegible]        STEROIDOGENIC ACUTE REGULATORY PROTEIN.
FT   CONFLICT   [illegible]        [illegible] (IN REF. [illegible]).
SQ   SEQUENCE   285 AA;  [illegible] MW;  [illegible]

  Query Match            [illegible]; Score [illegible]; DB [illegible]; Length 285;
  Best Local Similarity  [illegible]; Pred. No. [illegible]
  Matches    [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible]

Qy          1 NQEGWK 6

Db         [illegible] NQEGWK [illegible]


RESULT 7
TBE_HUMAN
ID   TBE_HUMAN      STANDARD;      PRT;   [illegible] AA.
AC   [illegible]
DT   [illegible] (Rel. [illegible], Created)
DT   [illegible] (Rel. [illegible], Last sequence update)
DT   [illegible] (Rel. [illegible], Last annotation update)
DE   TUBULIN EPSILON CHAIN (EPSILON TUBULIN).
GN   TUBE1 OR TUBE.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=2[illegible]; PubMed=[illegible]
RA   Chang P., Stearns T.;
RT   "[illegible] and [illegible] two new human centrosomal tubulins
RT   reveal new aspects of [illegible] structure and function.";
RL   Nat. [illegible]
CC   -!- SUBCELLULAR LOCATION: CENTROSOMAL [illegible] WITH [illegible]
CC       MATERIAL.
CC   -!- SIMILARITY: BELONGS TO THE TUBULIN FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use [illegible] content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
```

```
DR   EMBL; AF201334; AAF09585.1; -.
KW   Microtubules; GTP-binding.
FT   NP_BIND     148    154       GTP (POTENTIAL).
SQ   SEQUENCE   475 AA;  52931 MW;  3E8E71[illegible]BA6AFC80 CRC64;

  Query Match             100.0%;  Score 38;  DB 1;  Length 475;
  Best Local Similarity   100.0%;  Pred. No. 4.2;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 NQEGWK 6
              ||||||
Db        376 NQEGWK 381


RESULT   8
25A1_RAT
ID   25A1_RAT        STANDARD;      PRT;   358 AA.
AC   Q05961;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
     (2'-5')OLIGOADENYLATE SYNTHETASE 1 (EC 2.7.7.-) ((2-5')OLIGO(A)
     SYNTHETASE 1) (2-5A SYNTHETASE 1)
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93378704; PubMed=7764002;
RA   Truve E., Aaspollu A., Honkanen J., Puska R., Mehto M., Hassi A.,
RA   Teeri T.H., Kelve M., Seppanen P., Saarma M.;
RT   "Transgenic potato plants expressing mammalian 2'-5' oligoadenylate
RT   synthetase are protected from potato virus X infection under field
RT   conditions.";
RL   Biotechnology 11:1048-1052(1993).
CC   -!- CATALYTIC ACTIVITY: BINDS DOUBLE-STRANDED RNA AND POLYMERIZES ATP
CC       INTO PPP(A2'P5'A)N OLIGOMERS, WHICH ACTIVATE THE LATENT RNASE L
CC       THAT, WHEN ACTIVATED, CLEAVES SINGLE-STRANDED RNAS.
CC   -!- INDUCTION: BY INTERFERONS.
CC   -!- SIMILARITY: BELONGS TO THE 2-5A SYNTHETASE FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
     or send an email to license@isb-sib.ch).
     --------------------------------------------------------------------------
     EMBL; Z18877; CAA79317.1; -.
DR   PIR; S31407; S31407.
DR   InterPro; IPR001797; -.
DR   PROSITE; PS00832; 25A_SYNTH_1; 1
DR   PROSITE; PS00833; 25A_SYNTH_2; 1
KW   RNA-binding; Transferase; Nucleotidyltransferase;
KW   Interferon induction; Multigene family.
SQ   SEQUENCE   358 AA;  41666 MW;  A[illegible]0F[illegible]A5D45BEA CRC64;

  Query Match             92.1%;  Score 35;  DB 1;  Length 358;
  Best Local Similarity   83.3%;  Pred. No. 12;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 NQEGWK 6
              |||||:
Db        314 NQEGWR 319


RESULT   9
OAS1_RAT
ID   OAS1_RAT        STANDARD;      PRT;   358 AA.
AC   Q05961;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   2'-5'-OLIGOADENYLATE SYNTHETASE 1 (EC 2.7.7.-) ((2-5')OLIGO(A)
DE   SYNTHETASE 1) (2-5A SYNTHETASE 1).
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93378704; PubMed=7764002;
RA   Truve E., Aaspollu A., Honkanen J., Puska R., Mehto M., Hassi A.,
RA   Teeri T.H., Kelve M., Seppanen P., Saarma M.;
RT   "Transgenic potato plants expressing mammalian 2'-5' oligoadenylate
RT   synthetase are protected from potato virus X infection under field
RT   conditions.";
RL   Biotechnology 11:1048-1052(1993).
CC   -!- FUNCTION: MAY PLAY A ROLE IN MEDIATING RESISTANCE TO VIRUS
CC       INFECTION, CONTROL OF CELL GROWTH, DIFFERENTIATION, AND APOPTOSIS.
CC   -!- CATALYTIC ACTIVITY: BINDS DOUBLE-STRANDED RNA AND POLYMERIZES ATP
CC       INTO PPP(A2'P5'A)N OLIGOMERS, WHICH ACTIVATE THE LATENT RNASE L
CC       THAT, WHEN ACTIVATED, CLEAVES SINGLE-STRANDED RNAS.
CC   -!- INDUCTION: BY INTERFERONS.
CC   -!- SIMILARITY: BELONGS TO THE 2-5A SYNTHETASE FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; Z18877; CAA79317.1; -.
DR   PIR; S31407; S31407.
DR   InterPro; IPR001797; -.
DR   PROSITE; PS00832; 25A_SYNTH_1; 1.
DR   PROSITE; PS00833; 25A_SYNTH_2; 1.
DR   PROSITE; PS50152; 25A_SYNTH_3; 1.
KW   RNA-binding; Transferase; Nucleotidyltransferase;
KW   Interferon induction.
SQ   SEQUENCE   358 AA;  41666 MW;  AE0F[illegible]E6A7D45BEA CRC64;

  Query Match             92.1%;  Score 35;  DB 1;  Length 358;
  Best Local Similarity   83.3%;  Pred. No. 12;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 NQEGWK 6
              |||||:
Db        314 NQEGWR 319


RESULT  10
LYS1_SCHPO
ID   LYS1_SCHPO     STANDARD;      PRT;   [illegible]8 AA.
AC   Q09694; Q9U[illegible]C.
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-OCT-2000 (Rel. 40, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   SACCHAROPINE DEHYDROGENASE [NAD+, L-LYSINE FORMING] (EC 1.5.1.7)
DE   (LYSINE--2-OXOGLUTARATE REDUCTASE) (SDH).
GN   LYS3 OR SPAC227.18 OR SPAC2F7.01
OS   Schizosaccharomyces pombe (Fission yeast).
OC   Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC   Schizosaccharomycetales; Schizosaccharomycetaceae;
OC   Schizosaccharomyces.
OX   NCBI_TaxID=4896;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=972;
```

```
RA   Zimmermann W., Wambutt R., [illegible]
RA   Barrell B.G.;
RL   Submitted (NOV-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE OF 320-368 FROM N.A.
RC   STRAIN=972;
RA   Gentles S., [illegible] M., Barrell B.G., [illegible]
RL   Submitted (JUL-1995) to the EMBL/GenBank/DDBJ databases.
CC   -!- CATALYTIC ACTIVITY: N6-(L-1,3-DICARBOXYPROPYL)-L-LYSINE + NAD+ +
CC       H(2)O = L-LYSINE + 2-OXOGLUTARATE + NADH.
CC   -!- PATHWAY: 8TH STEP IN LYSINE BIOSYNTHESIS
CC   -!- SIMILARITY: TO OTHER FUNGAL SACCHAROPINE DEHYDROGENASE
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
     modified and this statement is not removed. Usage by and for commercial
     entities requires a license agreement (See [illegible]
     or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AL133156; CAB[illegible]
DR   EMBL; Z50142; CAA904[illegible]
KW   Lysine biosynthesis; Oxidoreductase; NAD.
FT   ACT_SITE    205    2[illegible]    BY SIMILARITY.
SQ   SEQUENCE   368 AA;  [illegible] MW;  [illegible] CRC64;

  Query Match             [illegible]%;  Score [illegible];  DB [illegible];  Length [illegible];
  Best Local Similarity   [illegible]%;  Pred. No. [illegible];
  Matches    5;  Conservative    [illegible];  Mismatches    [illegible];  Indels    [illegible];  Gaps    [illegible];

Qy          1 NQEGWK 6
              |||:|:
Db        100 NQEGWR 105


RESULT 11
STAR_BOVIN
ID   STAR_BOVIN     STANDARD;      PRT;   285 AA.
AC   Q28918;
DT   01-NOV-1997 (Rel. 35, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN PRECURSOR (
     STAR.
     Bos taurus (Bovine).
     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; [illegible]
OC   Mammalia; Eutheria; [illegible]
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Corpus luteum;
RX   MEDLINE=96011827; PubMed=7488[illegible]
RA   Hartung S., Rust W., [illegible] M., [illegible]
RT   "Molecular cloning and in vivo expression of the bovine steroidogenic
RT   acute regulatory protein.";
RL   Biochem. Biophys. Res. Commun. [illegible]
RN   [2]
RP   REVISIONS TO [illegible] AND [illegible]
RA   Hartung S., Rust W., [illegible] M., [illegible] R.;
RL   Submitted (MAY-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: PLAYS A KEY ROLE IN STEROID HORMONE SYNTHESIS BY
CC       ENHANCING THE METABOLISM OF CHOLESTEROL IN [illegible]
CC       TRANSPORTER THAT BINDS TO AND TRANSPORT THE ESTERS [illegible]
CC       INTERMEMBRANE SPACE OF THE MITOCHONDRION (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL.
CC   -!- TISSUE SPECIFICITY: CORPUS LUTEUM AND ADRENAL GLAND
CC   -!- SIMILARITY: BELONGS TO THE STAR FAMILY
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; Y[illegible]; CAA[illegible]
DR   EMBL; Y[illegible]; CAA[illegible]
DR   InterPro; IPR[illegible]
DR   InterPro; IPR[illegible]
DR   Pfam; PF[illegible]
DR   PRINTS; PR[illegible]; STARPROTEIN
KW   Lipid transport; Lipid transport; Transport; Steroidogenesis;
KW   Mitochondrion; Transit peptide.
FT   TRANSIT       1     [illegible]       MITOCHONDRION (POTENTIAL).
FT   CHAIN       [illegible]   285       STEROIDOGENIC ACUTE REGULATORY PROTEIN.
FT   CONFLICT    [illegible]   [illegible]       [illegible] -> Y (IN GENOMIC SEQUENCE).
FT   CONFLICT    [illegible]   [illegible]       [illegible] -> K (IN GENOMIC SEQUENCE).
SQ   SEQUENCE   285 AA;  [illegible] MW;  [illegible] CRC64;

  Query Match             [illegible]%;  Score [illegible];  DB [illegible];  Length 285;
  Best Local Similarity   [illegible]%;  Pred. No. [illegible];
  Matches    [illegible];  Conservative    [illegible];  Mismatches    [illegible];  Indels    0;  Gaps    0;

Qy          1 NQEGWK 6

Db         [illegible] NQEGWK [illegible]


RESULT 12
[illegible]
ID   [illegible]     STANDARD;      PRT;   [illegible] AA.
AC   P46[illegible];
DT   01-NOV-[illegible] (Rel. [illegible], Created)
DT   01-NOV-[illegible] (Rel. [illegible], Last sequence update)
DT   [illegible] (Rel. [illegible], Last annotation update)
DE   PUTATIVE EXOGLUCANASE TYPE C PRECURSOR (EC 3.2.1.91)
DE   (EXOCELLOBIOHYDROLASE C) (1,4-BETA-CELLOBIOHYDROLASE)
DE   (BETA-GLUCANCELLOBIOHYDROLASE)
OS   Fusarium oxysporum.
OC   Eukaryota; Fungi; Ascomycota; Pezizomycotina; [illegible]; Fusarium.
OX   NCBI_TaxID=[illegible]
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=[illegible]; PubMed=[illegible]
RA   Sheppard P.O., [illegible] A., Foster D.C.,
RA   Hagen F.S., Upshall A., McKnight G.L., [illegible]
RT   "The use of conserved cellulase family-specific sequences to clone
RT   cellulase homologue cDNAs from Fusarium oxysporum.";
RL   Gene [illegible]
CC   -!- CATALYTIC ACTIVITY: HYDROLYSIS OF 1,4-BETA-D-GLUCOSIDIC
CC       LINKAGES IN CELLULOSE AND CELLOTETRAOSE, RELEASING CELLOBIOSE
CC       FROM THE NON-REDUCING ENDS OF THE CHAINS.
CC   -!- SIMILARITY: CONTAINS 1 FUNGAL-TYPE CELLULOSE-BINDING DOMAIN (CBD).
CC   -!- SIMILARITY: BELONGS TO CELLULASE FAMILY 7 (FAMILY 7 OF GLYCOSYL
CC       HYDROLASES).
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; [illegible]; AAA[illegible]
DR   PDB; [illegible]
DR   InterPro; IPR[illegible]
DR   InterPro; IPR[illegible]
DR   Pfam; PF[illegible]
DR   Pfam; PF[illegible]
DR   PRINTS; PR[illegible]; GLHYDRLASE7
DR   PROSITE; PS00562; CBD_FUNGAL; 1.
KW   Cellulose degradation; Hydrolase; Glycosidase; Glycoprotein; Signal.
FT   SIGNAL        1     17         POTENTIAL.
FT   CHAIN        18    514         PUTATIVE EXOGLUCANASE TYPE C.
FT   DOMAIN       18    439         CATALYTIC
FT   DOMAIN      440    482         LINKER.
FT   DOMAIN      483    514         CELLULOSE BINDING (BY SIMILARITY).
FT   ACT_SITE    229    229         NUCLEOPHILE (BY SIMILARITY).
FT   ACT_SITE    234    234         PROTON DONOR (BY SIMILARITY).
FT   DISULFID    486    503         BY SIMILARITY.
FT   DISULFID    497    513         BY SIMILARITY.
FT   CARBOHYD    287    287         N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    490    490         N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   514 AA;  54704 MW;  6A4617[illegible]A46E[illegible]62 CRC64;

  Query Match             86.8%;  Score [illegible];  DB 1;  Length 514;
  Best Local Similarity   83.3%;  Pred. No. [illegible];
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 NQEGWK 6
              | ||||
          205 NSEGWK 210


RESULT  13
YG18_YEAST
ID   YG18_YEAST     STANDARD;      PRT;   620 AA.
AC   P53207;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   HYPOTHETICAL 71.4 KDA PROTEIN IN SEC9-MSB2 INTERGENIC REGION.
GN   YGR013W.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288C;
RX   MEDLINE=97435481; PubMed=9290212;
RA   Rieger M., Brueckner M., Schaefer M., Mueller-Auer S.;
RT   "Sequence analysis of 203 kilobases from Saccharomyces cerevisiae
RT   chromosome VII.";
RL   Yeast 13:1077-1090(1997).
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; Z72798; CAA96996.1; -.
DR   SGD; S0003245; SNU71.
KW   Hypothetical protein.
SQ   SEQUENCE   620 AA;  71382 MW;  2659A[illegible]2C4993D5 CRC64;

  Query Match             86.8%;  Score [illegible];  DB 1;  Length 620;
  Best Local Similarity   83.3%;  Pred. No. 48;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 NQEGWK 6
              :|.|||
Db         15 SQEGWK 20


RESULT  14
AMT3_CAEEL
ID   AMT3_CAEEL     STANDARD;      PRT;   622 AA.
AC   Q21565;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   PUTATIVE AMMONIUM TRANSPORTER M195.3.
GN   M195.3.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   Burton J.;
RL   Submitted (OCT-1995) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: INVOLVED IN THE UPTAKE OF AMMONIA (PROBABLE).
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (PROBABLE).
CC   -!- SIMILARITY: BELONGS TO THE AMT/MEP/NRGA FAMILY OF AMMONIUM
CC       TRANSPORTERS (TC 2.49).
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; Z66498; CAA91293.1; -.
DR   WormPep; M195.3; CE03516.
DR   InterPro; IPR001905; -.
DR   Pfam; PF00909; Ammonium_transp; 1.
DR   PROSITE; PS01219; AMMONIUM_TRANSP; 1.
KW   Hypothetical protein; Transport; Transmembrane.
FT   TRANSMEM     39     59         POTENTIAL.
FT   TRANSMEM     77     97         POTENTIAL.
FT   TRANSMEM    134    154         POTENTIAL.
FT   TRANSMEM    162    182         POTENTIAL.
FT   TRANSMEM    196    216         POTENTIAL.
FT   TRANSMEM    240    260         POTENTIAL.
FT   TRANSMEM    272    292         POTENTIAL.
FT   TRANSMEM    299    319         POTENTIAL.
FT   TRANSMEM    323    343         POTENTIAL.
FT   TRANSMEM    352    372         POTENTIAL.
FT   TRANSMEM    404    424         POTENTIAL.
FT   TRANSMEM    571    591         POTENTIAL.
SQ   SEQUENCE   622 AA;  67384 MW;  D6[illegible]BE CRC64;

  Query Match             86.8%;  Score [illegible];  DB 1;  Length 622;
  Best Local Similarity   100.0%;  Pred. No. 48;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 NQEGW 5
              |||||
Db        127 NQEGW 131


RESULT  15
SRF1_BACSU
ID   SRF1_BACSU     STANDARD;      PRT;   3587 AA.
AC   P27206;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   SURFACTIN SYNTHETASE SUBUNIT 1.
GN   SRFAA OR SRFA1 OR SRFA.
OS   Bacillus subtilis.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group;
OC   Bacillus/Staphylococcus group; Bacillus.
OX   NCBI_TaxID=1423;
RN   [1]
RP   SEQUENCE FROM N.A.
```

```
RC   STRAIN 168;
RX   MEDLINE 93181190; PubMed 8441[illegible];
RA   Fuma S., Fujishima Y., Corbell N., [illegible] Nakano M.M.,
RA   Zuber P., Yamane K.
RT   "Nucleotide sequence of 5' portion of srfA [illegible]
RT   required for competence establishment in Bacillus subtilis.";
RL   Nucleic Acids Res. [illegible]
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN 168 / JH642;
RX   MEDLINE 93360813; PubMed 835[illegible];
RA   Cosmina P., Rodriguez F., de Ferra F., Grandi G., [illegible]
RA   Venema G., van Sinderen D.;
RT   "Sequence and analysis of the genetic [illegible]
RT   synthesis in Bacillus subtilis.";
RL   Mol. Microbiol. 8:821-831(1993).
RN   [3]
[illegible]   SEQUENCE FROM N.A.
[illegible]   STRAIN 168;
[illegible]   MEDLINE 97124[illegible]; PubMed [illegible];
RA   Yamane K., Kumano M., Kurita K.;
RT   "The 25 degrees-36 degrees region of the [illegible]
RT   chromosome: determination of the sequence of a [illegible] and
RT   identification of [illegible] genes.";
RL   Microbiology 14[illegible]:3047-3056(1996).
RN   [4]
RP   SEQUENCE OF 1-46 FROM N.A.
RX   MEDLINE 91154131; PubMed 184[illegible];
RA   Nakano M.M., Magnuson R., Myers A.M., Curry [illegible]
RA   Zuber P.;
RT   "srfA is an operon required for surfactin production, competence
RT   development, and efficient sporulation [illegible]
RL   J. Bacteriol. 173:1770-1778(1991).
RN   [5]
RP   SEQUENCE OF 1-39 FROM N.A.
RX   MEDLINE 91358[illegible]; PubMed [illegible];
RA   Nakano M.M., Xia L., Zuber P.;
RT   "Transcription initiation region of the srfA [illegible]
RT   controlled by the comP-comA signal transduction [illegible]
RT   subtilis.";
RL   J. Bacteriol. 17[illegible]:5487-5493(1991).
RN   [6]
RP   SEQUENCE OF 1-64 FROM N.A.
RC   STRAIN 168;
RX   MEDLINE 95219[illegible]; PubMed [illegible];
[illegible]   Fujishima Y., Yamane K.;
[illegible]   "A 10 kb nucleotide sequence at the 5' [illegible]
[illegible]   of srfAA of the Bacillus subtilis chromosome.";
RL   Microbiology 141:277-279(1995).
CC   -!- FUNCTION: THIS PROTEIN IS A MULTIFUNCTIONAL ENZYME [illegible]
CC       ACTIVATE AND POLYMERIZE THE AMINO ACIDS [illegible]
CC       ACTIVATION SITES FOR THESE AA [illegible] DOMAIN.
CC   -!- COFACTOR: CONTAINS 3 COVALENTLY BOUND PHOSPHOPANTETHEINES.
CC   -!- PATHWAY: NON RIBOSOMAL BIOSYNTHESIS OF THE [illegible]
CC       ANTIBIOTIC SURFACTIN.
CC   -!- SIMILARITY: [illegible] OTHER ENZYMES WHICH [illegible] VIA AN [illegible]
CC       COVALENT BINDING OF AMP TO THEIR SUBSTRATE.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. [illegible]
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. [illegible]
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. [illegible]
CC   entities requires a license agreement (See http://[illegible]
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; D13262; BAA02525.1; -.
DR   EMBL; X70356; CAA49810.1; -.
DR   EMBL; D50453; BAA08981.1; -.
DR   EMBL; M59939; AAA22810.1; -.
DR   EMBL; M64702; AAA22810.1; -.
DR   EMBL; D36762; BAA07034.1; -.
DR   EMBL; Z[illegible]; CAB12514.1; -.
```

```
DR   PIR; [illegible]
DR   SUBTILIST; BG[illegible]; srfAA.
DR   InterPro; IPR[illegible];
DR   InterPro; IPR[illegible];
DR   InterPro; IPR[illegible];
DR   Pfam; PF00501; AMP-binding; 3.
DR   Pfam; PF[illegible]; [illegible]; 3.
DR   Pfam; PF00550; pp-binding; 3.
DR   PRINTS; [illegible]; AMPBINDING.
DR   PROSITE; PS00012; PHOSPHOPANTETHEINE; [illegible].
DR   PROSITE; PS00455; AMP_BINDING; [illegible].
DR   PROSITE; PS50075; ACP_DOMAIN; 3.
KW   Ligase; Antibiotic biosynthesis; Phosphopantetheine; Sporulation;
KW   Multifunctional enzyme; Repeat.
FT   REPEAT      [illegible]  [illegible]    DOMAIN 1 (GLU ACTIVATING).
FT   REPEAT      [illegible]  [illegible]    DOMAIN 2 (LEU ACTIVATING).
FT   REPEAT      [illegible]  [illegible]    DOMAIN 3 (LEU ACTIVATING).
FT   DOMAIN      [illegible]  [illegible]    ACYL CARRIER (ACP).
FT   DOMAIN      [illegible]  [illegible]    ACYL CARRIER (ACP).
FT   DOMAIN      [illegible]  [illegible]    ACYL CARRIER (ACP).
FT   BINDING     [illegible]  [illegible]    PHOSPHOPANTETHEINE (POTENTIAL).
FT   BINDING     2045   2045       PHOSPHOPANTETHEINE (POTENTIAL).
FT   BINDING     [illegible]  [illegible]    PHOSPHOPANTETHEINE (POTENTIAL).
FT   CONFLICT    149    149        [illegible] -> M (IN REF. [illegible] AND 3).
FT   CONFLICT    [illegible]  [illegible]    [illegible] -> V (IN REF. [illegible] AND 3).
FT   CONFLICT    16[illegible]  165  A -> G (IN REF. [illegible] AND 4).
FT   CONFLICT    [illegible]  [illegible]    [illegible] -> L (IN REF. [illegible] AND 3).
FT   CONFLICT    4[illegible]  [illegible]    [illegible] -> I (IN REF. [illegible] AND 3).
FT   CONFLICT    [illegible]  [illegible]    [illegible] -> A (IN REF. [illegible] AND 3).
FT   CONFLICT    [illegible]  [illegible]    [illegible] -> Y (IN REF. [illegible] AND 3).
FT   CONFLICT    [illegible]  [illegible]    [illegible] (IN REF. [illegible] AND 3).
FT   CONFLICT    641    641        [illegible] -> I (IN REF. [illegible] AND 3).
FT   CONFLICT    [illegible]  [illegible]    [illegible] -> F (IN REF. [illegible] AND 3).
FT   CONFLICT    [illegible]  [illegible]    A -> F (IN REF. [illegible] AND 3).
FT   CONFLICT    [illegible]  [illegible]    [illegible] -> L (IN REF. [illegible] AND 3).
FT   CONFLICT    [illegible]  [illegible]    [illegible]AVVNMLAAA[illegible]KLFRAF
FT                                [illegible]RRHESLF[illegible]RWRGSHEESGASSEMERSAIHASSSN
FT                                [illegible]RYSRAGAR[illegible]NLLERVIF (IN REF. [illegible] AND
FT                                [illegible]
FT   CONFLICT    [illegible]  [illegible]    [illegible] -> V[illegible] (IN REF. [illegible] AND 3).
FT   CONFLICT    [illegible]  [illegible]    [illegible] -> A (IN REF. [illegible] AND 3).
FT   CONFLICT    [illegible]  [illegible]    [illegible] (IN REF. [illegible] AND 3).
FT   CONFLICT    [illegible]  [illegible]    [illegible] -> I (IN REF. [illegible] AND 3).
FT   CONFLICT    [illegible]  [illegible]    [illegible] -> I (IN REF. [illegible] AND 3).
FT   CONFLICT    [illegible]  [illegible]    [illegible] -> A (IN REF. [illegible] AND 3).
FT   CONFLICT    [illegible]  [illegible]    [illegible]SKW[illegible]AV[illegible]M (IN REF. [illegible] AND 3).
FT   CONFLICT    [illegible]  [illegible]    [illegible] -> I (IN REF. [illegible] AND 3).
FT   CONFLICT    [illegible]  [illegible]    [illegible] -> L (IN REF. [illegible] AND 3).
FT   CONFLICT    [illegible]  [illegible]    [illegible] -> HE (IN REF. [illegible] AND 3).
FT   CONFLICT    [illegible]  [illegible]    V -> [illegible] (IN REF. [illegible] AND 3).
FT   CONFLICT    [illegible]  [illegible]    [illegible] (IN REF. [illegible] AND 3).
FT   CONFLICT    2[illegible]  [illegible]    [illegible] -> S (IN REF. [illegible] AND 3).
FT   CONFLICT    [illegible]  [illegible]    R[illegible] -> AV (IN REF. 2 AND 3).
FT   CONFLICT    [illegible]  [illegible]    [illegible] -> ENEL (IN REF. [illegible] AND 3).
FT   CONFLICT    [illegible]  [illegible]    [illegible] -> S (IN REF. 2 AND 3).
FT   CONFLICT    [illegible]  [illegible]    [illegible] -> A (IN REF. [illegible] AND 3).
FT   CONFLICT    [illegible]  [illegible]    [illegible] -> N (IN REF. 2 AND 3).
FT   CONFLICT    [illegible]  [illegible]    [illegible] -> N (IN REF. [illegible] AND 3).
FT   CONFLICT    [illegible]  [illegible]    A -> S (IN REF. 2 AND 3).
FT   CONFLICT    [illegible]  [illegible]    R -> S (IN REF. [illegible] AND 3).
FT   CONFLICT    3451   [illegible]    Y -> S (IN REF. 2 AND 3).
FT   CONFLICT    [illegible]  [illegible]    [illegible] -> RL (IN REF. [illegible] AND 3).
FT   CONFLICT    [illegible]  [illegible]    [illegible]ACL -> HPPF (IN REF. 2 AND 3).
FT   CONFLICT    [illegible]  [illegible]    [illegible] -> L (IN REF. [illegible] AND 3).
FT   CONFLICT    [illegible]  [illegible]    [illegible] -> H (IN REF. 2 AND 3).
FT   CONFLICT    [illegible]  [illegible]    [illegible] -> FH (IN REF. 2 AND 3).
FT   CONFLICT    [illegible]  [illegible]    [illegible] -> E (IN REF. [illegible] AND 3).
SQ   SEQUENCE   [illegible] AA;  [illegible] MW;  [illegible] CRC64;

  Query Match              [illegible]
  Best Local Similarity    [illegible]  Pred. No. [illegible]
  Matches     5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps     ;

Qy       1 NQEGWK 6
           || |||
Db    3084 NQHGWK 3089


Search completed: August 22, 2001, 14:30:57
Job time: 411 sec
```

```
                     GenCore version 1.5
           Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on:          August 22, 2001, 14:27:25  Search time 112.12 Seconds
                                            (without alignments)
                                            7.080 Million cell updates/sec

Title:           US-08-659-254-6
Perfect score:   38
Sequence:        1 NQEGWK 6

Scoring table:   BLOSUM62
                 Gapop 10.0  Gapext 0.5

Searched:        425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters:      425026

   mum DB seq length: 0
   mum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       SPTREMBL_16:*
                1:  sp_archea:*
                2:  sp_bacteria:*
                3:  sp_fungi:*
                4:  sp_human:*
                5:  sp_invertebrate:*
                6:  sp_mammal:*
                7:  sp_mhc:*
                8:  sp_organelle:*
                9:  sp_phage:*
                10:  sp_plant:*
                11:  sp_rodent:*
                12:  sp_unclassified:*
                13:  sp_vertebrate:*
                14:  sp_virus:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                              SUMMARIES

            %
  lt      Query
 No.  Score Match Length DB  ID                          Description
------------------------------------------------------------------------
   1    35  92.1    102  4  Q9H214                   Q9h214 homo sapien
   2    35  92.1    783  2  Q45554                   Q45554 bacillus sp
   3    35  92.1    824  2  Q9F216                   Q9f216 bacillus sp
   4    34  89.5    449 10   Q9FMJ8                   Q9fmj8 arabidopsis
   5    33  86.8    207  2  O32813                   O32813 lactococcus
   6    33  86.8    244  5  Q9U417                   Q9u417 branchiosto
   7    33  86.8    281 13   Q9DQ09                   Q9dq09 gallus gall
   8    33  86.8    285  6  Q9GMD0                   Q9gmd0 ovis aries
   9    33  86.8    347  5  O61700                   O61700 branchiosto
  10    33  86.8    369  5  O96622                   O96622 dictyosteli
  11    33  86.8    433  2  Q9KBW9                   Q9kbw9 bacillus ha
  12    33  86.8    472 13   O93342                   O93342 gallus gall
  13    33  86.8    695  8  Q31773                   Q31773 atropa bell
  14    33  86.8   3583  2  Q45675                   Q45675 bacillus sa
  15    33  86.8   4564  5  O77075                   O77075 drosophila
  16    32  84.2    271  2  O25947                   O25947 helicobacte
  17    32  84.2    315  2  Q43888                   Q43888 streptococ
  18    32  84.2    329 14   Q9FIV8                   Q9fiv8 edge hill v
  19    32  84.2    329 14   Q9FIV5                   Q9fiv5 edge hill v
```

```
  20    32  84.2    333  6  Q9GL24                   Q9gl24 canis famil
  21    32  84.2    333  6  Q9GKI8                   Q9gki8 cercopithec
  22    32  84.2    374  1  Q50245                   Q50245 methanosarc
  23    32  84.2    398  3  Q05765                   Q05765 saccharomyc
  24    32  84.2    460  4  Q9Y589                   Q9y589 homo sapien
  25    32  84.2    460  4  Q9NWZ3                   Q9nwz3 homo sapien
  26    32  84.2    476  5  Q93513                   Q93513 caenorhabdi
  27    32  84.2    500 10   Q9SV26                   Q9sv26 arabidopsis
  28    32  84.2    616 14   Q8[illegible]52                   Q8[illegible]52 newcastle d
  29    32  84.2    714  2  O68005                   O68005 enterococcu
  30    32  84.2    714  2  Q60149                   Q60149 streptococc
  31    32  84.2    716  2  O34575                   O34575 bacillus su
  32    32  84.2    847  4  Q16584                   Q16584 homo sapien
  33    32  84.2    850 11   Q9JJ15                   Q9jj15 mus musculu
  34    32  84.2   1125  5  Q93203                   Q93203 caenorhabdi
  35    32  84.2   1943  2  O25331                   O25331 helicobacte
  36    32  84.2   2115 14   Q8650[illegible]                   Q8650[illegible] rubella vir
  37    32  84.2   2116 14   O40955                   O40955 rubella vir
  38    32  84.2   2129 14   Q9J6K9                   Q9j6k9 rubella vir
  39    31  81.6    423  3  Q9UVP4                   Q9uvp4 alternaria
  40    31  81.6    464 10   Q9LNY0                   Q9lny0 arabidopsis
  41    31  81.6    539  5  Q23305                   Q23305 caenorhabdi
  42    31  81.6    727  5  O45675                   O45675 caenorhabdi
  43    31  81.6   4167  5  Q9GPN8                   Q9gpn8 drosophila
  44    30  78.9    129  2  Q9I345                   Q9i345 pseudomonas
  45    30  78.9    131  2  Q9X[illegible]23                   Q9x[illegible]23 bacillus an

                              ALIGNMENTS

RESULT   1
Q9H214
ID   Q9H214      PRELIMINARY;      PRT;    102 AA.
AC   Q9H214;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   DC42.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Xu X., Yang Y., Gao G., Xiao H., Chen Z., Han Z.;
RL   Submitted (MAY-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF267864; AAG44733.1; -.
SQ   SEQUENCE   102 AA;  12247 MW;  A1[illegible]27A64ECD59 [illegible]64;

  Query Match             92.1%;  Score 35;  DB 4;  Length 102;
  Best Local Similarity   83.3%;  Pred. No. 19;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 NQEGWK 6
              ||:|||
Db         86 NQQGWK 91


RESULT   2
Q45554
ID   Q45554      PRELIMINARY;      PRT;    783 AA.
AC   Q45554;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   ENDOGLUCANASE PRECURSOR (EC 3.2.1.4) (ENDO-1,4-BETA-GLUCANASE)
DE   (ALKALINE CELLULASE).
OS   Bacillus SP.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group;
OC   Bacillus/Staphylococcus group; Bacillus.
OX   NCBI_TaxID=1409;
```

```
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN 22 28;
RA   Miyatake M., [illegible]
RL   Submitted (MAY 1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS [illegible]
CC       LINKAGES IN CELLULOSE (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO CELLULASE FAMILY A [illegible]
CC       HYDROLASES).
DR   EMBL; D85236; BAA12[illegible]
DR   HSSP; [illegible]
DR   InterPro; IPR00[illegible]
DR   Pfam; PF001[illegible]
DR   PROSITE; PS[illegible]
KW   Cellulose degradation; Hydrolase; [illegible]
FT   SIGNAL      [illegible]
FT   CHAIN       [illegible]       ALKALINE [illegible]
FT   ACT_SITE    [illegible]       [illegible]
FT   ACT_SITE    [illegible]       [illegible]
SQ   SEQUENCE   [illegible] AA;  [illegible] MW;  [illegible]

  Query Match             [illegible]
  Best Local Similarity   [illegible]
  Matches    5;  Conservative    1;  Mismatches  [illegible]

Qy          1 NQEGWK 6
               |||||
Db        204 NEEGWK 2[illegible]


RESULT 3
Q9EZ16
ID   Q9EZ16      PRELIMINARY;      PRT;   [illegible] AA.
AC   Q9EZ16;
DT   01 MAR 2001 (TrEMBLrel. 16, Created)
DT   01 MAR 2001 (TrEMBLrel. 16, Last sequence update)
DT   01 MAR 2001 (TrEMBLrel. 16, Last annotation update)
DE   CELLULASE.
OS   Bacillus sp.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group;
OC   Bacillus/Staphylococcus group; Bacillus.
OX   NCBI_TaxID=1409;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN KSM-S237;
RA   Hakamada Y., Hatada Y., Koike K., Yoshimatsu T., Kawai
RA   Kobayashi T., Ito S.;
RT   "Deduced amino acid sequence and possible [illegible]
RT   thermostable, alkaline cellulase from an alkaliphilic Bacillus
RT   strain.";
RL   Biosci. Biotechnol. Biochem. 64:2281-2289(2000).
DR   EMBL; AB018420; BAB19[illegible]
SQ   SEQUENCE   824 AA;  [illegible] MW;  [illegible]

  Query Match             [illegible]
  Best Local Similarity   [illegible]
  Matches    5;  Conservative    1;  Mismatches  [illegible]

Qy          1 NQEGWK 6
               |||||
Db        204 NEEGWK 2[illegible]


RESULT 4
Q9FEI8
ID   Q9FEI8      PRELIMINARY;      PRT;   [illegible] AA.
AC   Q9FEI8;
DT   01 MAR 2001 (TrEMBLrel. 16, Created)
DT   01 MAR 2001 (TrEMBLrel. 16, Last sequence update)
DT   01 MAR 2001 (TrEMBLrel. 16, Last annotation update)
DE   GENOMIC DNA, CHROMOSOME [illegible]
```

```
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; [illegible] Brassicales;
OC   Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=[illegible]
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN [illegible]
RX   MEDLINE=[illegible] PubMed=[illegible]
RA   [illegible] Nakamura Y., Kaneko T., [illegible] Tabata S.
RA   Miyajima N., [illegible]
RT   "[illegible] of Arabidopsis thaliana chromosome [illegible] sequence
RT   features of [illegible] physically assigned
RT   [illegible]";
RL   DNA Res. [illegible]
DR   EMBL; AB[illegible]; BAB0[illegible]
SQ   SEQUENCE   [illegible] AA;  [illegible] MW;  [illegible]

  Query Match             [illegible]  Length 44[illegible]
  Best Local Similarity   [illegible]  Pred. No. [illegible]
  Matches    [illegible]  Conservative  [illegible]  Mismatches  [illegible]  Indels  [illegible]  Gaps  [illegible]

Qy          1 NQ[illegible]WK [illegible]

Db        [illegible] NKEGWK [illegible]


RESULT [illegible]
[illegible]
ID   [illegible]      PRELIMINARY;      PRT;   [illegible] AA.
AC   [illegible]
DT   01 JAN [illegible] (TrEMBLrel. [illegible], Created)
DT   [illegible] JAN [illegible] (TrEMBLrel. [illegible], Last sequence update)
DT   [illegible] JUN [illegible] (TrEMBLrel. [illegible], Last annotation update)
DE   LACTOSE [illegible] LACTOSE [illegible] AND [illegible] GENES, PARTIAL CDS (FRAGMENT).
OS   [illegible] (subsp. cremoris) (Streptococcus cremoris).
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC   [illegible]
OX   NCBI_TaxID=[illegible]
RN   [illegible]
RP   SEQUENCE FROM N.A.
RC   STRAIN MG[illegible]
RX   MEDLINE=[illegible] PubMed=[illegible]
RA   [illegible]
RT   "Characterization of lactococcal [illegible] UV-sensitive mutants obtained
RT   by ISS1 transposition.";
RL   J. Bacteriol. [illegible]
DR   EMBL; [illegible] AA[illegible]
DR   InterPro; IPR00[illegible]
DR   InterPro; IPR00[illegible]
DR   InterPro; [illegible]
DR   Pfam; PF0[illegible] Ubie methyltransf; 1.
FT   NON_TER      [illegible]
SQ   SEQUENCE   [illegible] AA;  [illegible] MW;  [illegible]

  Query Match             [illegible]  Length 2[illegible]
  Best Local Similarity   [illegible]  Pred. No. [illegible]
  Matches    [illegible]  Conservative  [illegible]  Mismatches  [illegible]  Indels  [illegible]  Gaps  [illegible]

Qy          [illegible] N[illegible]WK [illegible]

Db        [illegible] N[illegible]WK [illegible]


RESULT [illegible]
[illegible]
ID   [illegible]      PRELIMINARY;      PRT;   [illegible] AA.
AC   [illegible]
DT   [illegible] MAY [illegible] (TrEMBLrel. [illegible], Created)
DT   [illegible] MAY [illegible] (TrEMBLrel. [illegible], Last sequence update)
DT   [illegible] MAR [illegible] (TrEMBLrel. [illegible], Last annotation update)
```

```
DE   SECRETED PROTEIN WNT7 (FRAGMENT).
GN   WNT7.
OS   Branchiostoma belcheri.
OC   Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;
OC   Branchiostoma.
OX   NCBI_TaxID=7741;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Yasui K., Saiga H., Uemura M., Semba [illegible]
RT   "Early body formation and expression pattern of genes encoding
RT   secreted proteins, BbBMP2/4, BbWnt7, and BbWnt8 in lancelets.";
RL   Submitted (NOV-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: PROBABLE DEVELOPMENTAL PROTEIN. MAY BE A SIGNALLING
CC       MOLECULE WHICH AFFECT THE DEVELOPMENT OF DISCRETE REGIONS OF
CC       TISSUES. IS LIKELY TO SIGNAL OVER ONLY FEW CELL DIAMETERS (BY
CC       SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: POSSIBLY SECRETED AND ASSOCIATES WITH THE
CC       EXTRACELLULAR MATRIX (BY SIMILARITY).
CC   -!- SIMILARITY: TO OTHER MEMBERS OF THE WNT FAMILY.
DR   EMBL; AF206499; AAF19839.1; -.
DR   InterPro; IPR000970; -.
DR   InterPro; IPR001781; -.
     Pfam; PF00110; wnt; 1.
     ProDom; PD000094; -; 1.
     PROSITE; PS00246; WNT1; 1.
DR   SMART; SM00097; WNT1; 1.
KW   Developmental protein; Glycoprotein
FT   NON_TER       1      1
SQ   SEQUENCE   244 AA;  27942 MW;  [illegible] CRC64;

  Query Match             86.8%;  Score [illegible];  DB 5;  Length 244;
  Best Local Similarity   83.3%;  Pred. No. [illegible];
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    [illegible];  Gaps    [illegible];

Qy         1 NQEGWK 6
             | ||||
Db        40 NDEGWK 45


RESULT    7
Q9DG09
ID   Q9DG09       PRELIMINARY;      PRT;   281 AA.
AC   Q9DG09;
DT   01-MAR-2001 (TrEMBLrel. 16, Created
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN
GN   STAR.
OS   Gallus gallus (Chicken).
     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
     Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
     Gallus
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   PubMed=11064158;
RA   Bauer M.P., Bridgham J.T., Langenau D.M., Johnson A.L., Goetz F.W.;
RT   "Conservation of steroidogenic acute regulatory (StAR) protein
RT   structure and expression in vertebrates.";
RL   Mol. Cell. Endocrinol. 168:119-125(2000).
DR   EMBL; AF220436; AAG28594.1; -.
SQ   SEQUENCE   281 AA;  31036 MW;  [illegible] CRC64.

  Query Match             86.8%;  Score [illegible];  DB 13;  Length 281;
  Best Local Similarity   83.3%;  Pred. No. 1.2e+02;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    [illegible];  Gaps    [illegible];

Qy         1 NQEGWK 6
             :|||||
Db        92 DQEGWK 97
```

```
RESULT    8
Q9GMD0
ID   Q9GMD0       PRELIMINARY;      PRT;   285 AA.
AC   Q9GMD0;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN.
GN   STAR.
OS   Ovis aries (Sheep).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Caprinae; Ovis.
OX   NCBI_TaxID=9940;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   West L.A., Horvat R.D., Roess D.A., Barisas B.G., Juengel J.L.,
RA   Niswender G.D.;
RT   "Steroidogenic Acute Regulatory Protein and Peripheral-Type
RT   Benzodiazepine Receptor Associate at the Mitochondrial Membrane.";
RL   Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF290202; AAG02464.1; -.
SQ   SEQUENCE   285 AA;  31895 MW;  [illegible] CRC64;

  Query Match             86.8%;  Score [illegible];  DB 6;  Length 285;
  Best Local Similarity   83.3%;  Pred. No. [illegible];
  Matches    5;  Conservative    1;  Mismatches    [illegible];  Indels    [illegible];  Gaps    0;

Qy         1 NQEGWK 6
             :||||
Db        92 DQEGWK 97


RESULT    9
O61700
ID   O61700       PRELIMINARY;      PRT;   347 AA.
AC   O61700;
DT   01-AUG-1998 (TrEMBLrel. 07, Created)
DT   01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   AMPHIWNT7B.
GN   WNT7B.
OS   Branchiostoma floridae (Florida lancelet) (Amphioxus).
OC   Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;
OC   Branchiostoma.
OX   NCBI_TaxID=7739;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Schubert M., Holland L.Z., Jacobs D.K., Holland N.D.;
RL   Submitted (APR-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: PROBABLE DEVELOPMENTAL PROTEIN. MAY BE A SIGNALLING
CC       MOLECULE WHICH AFFECT THE DEVELOPMENT OF DISCRETE REGIONS OF
CC       TISSUES. IS LIKELY TO SIGNAL OVER ONLY FEW CELL DIAMETERS (BY
CC       SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: POSSIBLY SECRETED AND ASSOCIATES WITH THE
CC       EXTRACELLULAR MATRIX.
CC   -!- SIMILARITY: TO OTHER MEMBERS OF THE WNT FAMILY.
DR   EMBL; AF061975; AAC80433.1; -.
DR   InterPro; IPR000970; -.
DR   Pfam; PF00110; wnt; 1.
DR   PRINTS; PR01349; WNTPROTEIN.
DR   PROSITE; PS00246; WNT1; 1.
DR   SMART; SM00097; WNT1; 1.
KW   Developmental protein; Glycoprotein.
SQ   SEQUENCE   347 AA;  39425 MW;  [illegible] CRC64;

  Query Match             86.8%;  Score [illegible];  DB 5;  Length 347;
  Best Local Similarity   83.3%;  Pred. No. [illegible];
  Matches    5;  Conservative    [illegible];  Mismatches    [illegible];  Indels    0;  Gaps    0;
```

```
Qy          1 NQEGWK 6
              | ||||
Db        140 NDEGWK 1[illegible]


RESULT 10
O96622
ID   O96622      PRELIMINARY;      PRT;   [illegible] AA.
AC   O96622;
DT   01 MAY 1999 (TrEMBLrel. 10, Created)
DT   01 MAY 1999 (TrEMBLrel. 10, Last sequence update)
DT   01 MAR 2001 (TrEMBLrel. 16, Last annotation update)
DE   P41 ARC.
GN   ARPD.
OS   Dictyostelium discoideum (Slime mold).
OC   Eukaryota; Mycetozoa; Dictyosteliida; [illegible]
[illegible]   NCBI_TaxID=44[illegible];
[illegible]   [1]
[illegible]   SEQUENCE FROM N.A.
RC   STRAIN=AX3;
RA   Insall R.H., Machesky L.M.;
RT   "The members of the Dictyostelium Arp2/Arp3 complex.";
RL   Submitted (OCT-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF095930; AAC[illegible]; -.
DR   InterPro; IPR001680; [illegible]
DR   Pfam; PF00400; WD40; 4.
DR   SMART; SM00320; WD40; [illegible]
KW   Repeat; WD repeat.
SQ   SEQUENCE   369 AA;  [illegible] MW;  [illegible] CRC64;

  Query Match             8[illegible]%;  Score [illegible];  DB [illegible];  Length [illegible];
  Best Local Similarity   8[illegible]%;  Pred. No. [illegible];
  Matches    5;  Conservative    [illegible];  Mismatches    [illegible];  Indels [illegible];  Gaps [illegible];

Qy          1 NQEGWK 6
              || |||
Db        273 NQNGWK 278


RESULT 11
Q9KBW9
ID   Q9KBW9      PRELIMINARY;      PRT;   [illegible] AA.
AC   Q9KBW9;
DT   01 OCT 2000 (TrEMBLrel. 15, Created)
[illegible]   01 OCT 2000 (TrEMBLrel. 15, Last sequence update)
[illegible]   01 MAR 2001 (TrEMBLrel. 16, Last annotation update)
DE   BH1805 PROTEIN.
GN   BH1805.
OS   Bacillus halodurans.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; [illegible]
OC   Bacillus/Staphylococcus group; Bacillus.
OX   NCBI_TaxID=86665;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C-125 / JCM 9153;
RA   Takami H., Nakasone K., Takaki Y.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AP001513; BAB0[illegible]; -.
DR   InterPro; IPR000761; -.
DR   PROSITE; PS00022; EGF_1; UNKNOWN_1.
SQ   SEQUENCE   433 AA;  4[illegible] MW;  [illegible] CRC64;

  Query Match             [illegible]%;  Score [illegible];  DB [illegible];  Length [illegible];
  Best Local Similarity   [illegible]%;  Pred. No. [illegible];
  Matches    5;  Conservative    [illegible];  Mismatches    [illegible];  Indels [illegible];  Gaps [illegible];

Qy          1 NQEGW 5
              |||||
Db        375 NQEGW 379


RESULT [illegible]
[illegible]
ID   [illegible]      PRELIMINARY;      PRT;   47[illegible] AA.
AC   [illegible]
DT   [illegible] NOV 1998 (TrEMBLrel. [illegible], Created)
DT   [illegible] NOV 1998 (TrEMBLrel. [illegible], Last sequence update)
DT   [illegible] MAR 2001 (TrEMBLrel. 16, Last annotation update)
DE   MATRIX METALLOPROTEINASE
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=[illegible];
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Yang M., Kurkinen M.;
RT   "[illegible] matrix metalloproteinase (MMP) from chicken
RT   embryo [illegible]."
RL   [illegible]
DR   EMBL; AF[illegible]; AA[illegible]; -.
DR   HSSP; [illegible]4; [illegible]
DR   MEROPS; M10[illegible];
DR   InterPro; IPR[illegible];
DR   InterPro; IPR[illegible];
DR   InterPro; IPR[illegible];
DR   Pfam; PF[illegible]; Hemopexin; 4.
DR   Pfam; PF[illegible]; Peptidase_M10; 1.
DR   PRINTS; PR00138; MATRIXIN.
DR   PROSITE; PS00546; CYSTEINE_SWITCH; [illegible]
DR   PROSITE; PS00024; HEMOPEXIN; [illegible]
DR   PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_[illegible].
DR   SMART; SM[illegible]; [illegible]
SQ   SEQUENCE   47[illegible] AA;  [illegible] MW;  AA[illegible] CRC64;

  Query Match             8[illegible]%;  Score [illegible];  DB [illegible];  Length 47[illegible];
  Best Local Similarity   8[illegible]%;  Pred. No. [illegible];
  Matches    [illegible];  Conservative    [illegible];  Mismatches    [illegible];  Indels    [illegible];  Gaps    [illegible];

Qy          [illegible] N[illegible]GWK [illegible]

Db        [illegible]4 N[illegible]GWK [illegible]


RESULT [illegible]
[illegible]
ID   [illegible]      PRELIMINARY;      PRT;   [illegible] AA.
AC   [illegible]
DT   [illegible] NOV [illegible] (TrEMBLrel. [illegible], Created)
DT   [illegible] MAR [illegible] (TrEMBLrel. 16, Last sequence update)
DT   [illegible] MAR [illegible] (TrEMBLrel. [illegible], Last annotation update)
DE   NADH DEHYDROGENASE SUBUNIT (FRAGMENT).
GN   NDHF.
OS   [illegible]
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; [illegible] Asteridae; euasterids I;
OC   Solanales; Solanaceae; [illegible]
OX   NCBI_TaxID=[illegible];
RN   [1]
RP   SEQUENCE FROM N.A.
RA   [illegible] K.M., [illegible]
RT   "A [illegible] the Asteridae sensu lato based on rbcL
RT   sequences."
RL   Ann. M[illegible]
RN   [2]
RP   SEQUENCE FROM N.A.
RA   [illegible], Sweere J.A.;
RT   "[illegible] data in phylogenetic systematics: an empirical approach
RT   [illegible] in the Solanaceae.";
RL   [illegible]
RN   [illegible]
RP   SEQUENCE FROM N.A.
```

```
RA   Olmstead R.G.;
RL   Submitted (APR-1994) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U08915; AAA18597.2; -.
DR   InterPro; IPR001516; -.
DR   InterPro; IPR001750; -.
DR   InterPro; IPR002128; -.
DR   Pfam; PF00361; oxidored_q1; 1.
DR   Pfam; PF00662; oxidored_q1_N; 1.
DR   Pfam; PF01010; oxidored_q1_C; 1.
KW   Chloroplast.
FT   NON_TER       1      1
FT   NON_TER     695    695
SQ   SEQUENCE   695 AA;  78723 MW;  62F96A[illegible] [illegible]0D[illegible]5E7 CRC64;


  Query Match             86.8%;  Score [illegible];  DB 8;  Length 695;
  Best Local Similarity   100.0%;  Pred. No. [illegible]e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 NQEGW 5
            |||||
Db      556 NQEGW 560


RESULT  14
Q45675
ID   Q45675      PRELIMINARY;      PRT;  [illegible] AA.
AC   Q45675;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   SURFACTIN SYNTHETASE.
OS   Bacillus subtilis.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group;
OC   Bacillus/Staphylococcus group; Bacillus
OX   NCBI_TaxID=1423;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=168 TRPC2;
RX   MEDLINE=95219089; PubMed=7704264;
RA   Fabret C., Quentin Y., Guiseppi A., Busuttil J., Haiech J.,
RA   Denizot F.;
RT   "Analysis of errors in finished DNA sequences: the surfactin operon of
RT   Bacillus subtilis as an example.";
RL   Microbiology 141:345-350(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=168 TRPC2;
RX   MEDLINE=93181186; PubMed=8441623;
RA   Fuma S., Fujishima Y., Corbell N., D'Souza C., Nakano M.M., Zuber P.,
     Yamane K.;
     "Nucleotide sequence of 5' portion of srfA that contains the region
     required for competence establishment in Bacillus subtilis.";
RL   Nucleic Acids Res. 21:93-97(1993).
DR   EMBL; X72672; CAA51223.1; -.
DR   HSSP; P14687; 1AMU.
DR   InterPro; IPR000255; -.
DR   InterPro; IPR000873; -.
DR   InterPro; IPR001242; -.
DR   Pfam; PF00501; AMP-binding; 3
DR   Pfam; PF00550; pp-binding; 3
DR   Pfam; PF00668; Condensation; 4.
DR   PRINTS; PR00154; AMPBINDING.
DR   PROSITE; PS50075; ACP_DOMAIN; 3.
DR   PROSITE; PS00455; AMP_BINDING; 3.
DR   PROSITE; PS00012; PHOSPHOPANTETHEINE; UNKNOWN_3.
KW   Phosphopantetheine.
FT   CONFLICT    113    113       R -> A (IN REF. 2).
FT   CONFLICT    940    940       L -> R (IN REF. 2).
FT   CONFLICT   1310   1310       V -> C (IN REF. 2).
FT   CONFLICT   1782   1782       T -> S (IN REF. 2).
FT   CONFLICT   1817   1817       G -> R (IN REF. 2).
FT   CONFLICT   2070   2070       R -> C (IN REF. 2).
```

```
FT   CONFLICT   2135   2135       R -> A (IN REF. 2).
FT   CONFLICT   2390   2390       L -> A (IN REF. 2).
FT   CONFLICT   2481   2481       A -> [illegible] (IN REF. 2).
FT   CONFLICT   2486   2486       A -> [illegible] (IN REF. 2).
FT   CONFLICT   2542   2542       T -> I (IN REF. 2).
FT   CONFLICT   2544   2544       H -> A (IN REF. 2).
FT   CONFLICT   2563   2563       N -> [illegible] (IN REF. 2).
FT   CONFLICT   2604   2604       EM -> [illegible]K (IN REF. 2)
FT   CONFLICT   2641   2641       L -> [illegible] (IN REF. 2).
FT   CONFLICT   2895   2895       R -> [illegible] (IN REF. 2).
SQ   SEQUENCE   3583 AA;  400941 MW;  A[illegible]7AC[illegible]4[illegible]4[illegible]64C[illegible] CRC64;

  Query Match             86.8%;  Score [illegible];  DB 2;  Length 3583;
  Best Local Similarity   83.3%;  Pred. No. [illegible]e+04;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        1 NQEGWK 6
            || |||
Db     3080 NQHGWK 3085


RESULT  15
O77075
ID   O77075      PRELIMINARY;      PRT;  4564 AA.
AC   O77075;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   DHC7.
GN   THREADS.
OS   Drosophila hydei (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7224;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=TESTES;
RX   MEDLINE=98315100; PubMed=9649526;
RA   Kurek R., Reugels A.M., Glatzer K.H., Bunemann H.;
RT   "The Y chromosomal fertility factor threads in Drosophila hydei
RT   harbors a functional gene encoding an axonemal dynein beta heavy chain
RT   protein.";
RL   Genetics 149:1363-1376(1998).
DR   EMBL; AF031494; AAC35745.1; -.
DR   FlyBase; FBgn0020731; Dhyd\threads.
DR   InterPro; IPR000169; -.
DR   InterPro; IPR000566; -.
DR   InterPro; IPR001993; -.
DR   InterPro; IPR003593; -.
DR   PROSITE; PS00213; LIPOCALIN; UNKNOWN_1.
DR   PROSITE; PS00215; MITOCH_CARRIER; UNKNOWN_1.
DR   PROSITE; PS00639; THIOL_PROTEASE_HIS; UNKNOWN_1.
DR   SMART; SM00382; AAA; 1.
SQ   SEQUENCE   4564 AA;  526553 MW;  [illegible]A[illegible]98BAEB CRC64;


  Query Match             86.8%;  Score [illegible];  DB 5;  Length 4564;
  Best Local Similarity   100.0%;  Pred. No. [illegible];
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 NQEGW 5
            |||||
Db      135 NQEGW 139



Search completed: August 22, 2001, 14:[illegible]
Job time : 201 sec
```

```
                    GenCore version 4.5
           Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on:          August 22, 2001, 14:29:1[illegible]  Search time [illegible] Seconds
                                              (without alignments)
                                              5.351 Million cell updates/sec

Title:           US-08-659-254-7
Perfect score:   53
Sequence:        1 AEHGPICMV 9

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        412676 seqs, 69623988 residues

Total number of hits satisfying chosen parameters:      412676

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       A_Geneseq_0601:*
               1:  /SIDS8/gcgdata/geneseq/geneseqp/AA1980.DAT:*
               2:  /SIDS8/gcgdata/geneseq/geneseqp/AA1981.DAT:*
               3:  /SIDS8/gcgdata/geneseq/geneseqp/AA1982.DAT:*
               4:  /SIDS8/gcgdata/geneseq/geneseqp/AA1983.DAT:*
               5:  /SIDS8/gcgdata/geneseq/geneseqp/AA1984.DAT:*
               6:  /SIDS8/gcgdata/geneseq/geneseqp/AA1985.DAT:*
               7:  /SIDS8/gcgdata/geneseq/geneseqp/AA1986.DAT:*
               8:  /SIDS8/gcgdata/geneseq/geneseqp/AA1987.DAT:*
               9:  /SIDS8/gcgdata/geneseq/geneseqp/AA1988.DAT:*
               10:  /SIDS8/gcgdata/geneseq/geneseqp/AA1989.DAT:*
               11:  /SIDS8/gcgdata/geneseq/geneseqp/AA1990.DAT:*
               12:  /SIDS8/gcgdata/geneseq/geneseqp/AA1991.DAT:*
               13:  /SIDS8/gcgdata/geneseq/geneseqp/AA1992.DAT:*
               14:  /SIDS8/gcgdata/geneseq/geneseqp/AA1993.DAT:*
               15:  /SIDS8/gcgdata/geneseq/geneseqp/AA1994.DAT:*
               16:  /SIDS8/gcgdata/geneseq/geneseqp/AA1995.DAT:*
               17:  /SIDS8/gcgdata/geneseq/geneseqp/AA1996.DAT:*
               18:  /SIDS8/gcgdata/geneseq/geneseqp/AA1997.DAT:*
               19:  /SIDS8/gcgdata/geneseq/geneseqp/AA1998.DAT:*
               20:  /SIDS8/gcgdata/geneseq/geneseqp/AA1999.DAT:*
               21:  /SIDS8/gcgdata/geneseq/geneseqp/AA2000.DAT:*
               22:  /SIDS8/gcgdata/geneseq/geneseqp/AA2001.DAT:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                 SUMMARIES

                   %
Result            Query
  No.    Score    Match Length DB  ID                         Description
-----------------------------------------------------------------------------
     1      53    100.0      9  22  AAY72836                  [illegible]
     2      53    100.0     14  20  AAW74153                  Mouse steroidogeni
     3      53    100.0     14  22  AAY72835                  Mouse tryptic pept
     4      53    100.0    284  17  AAW13238                  Human steroidogene
     5      53    100.0    284  20  AAW74154                  Mouse steroidogeni
     6      53    100.0    284  22  AAY72834                  Mouse steroidogeni
     7      53    100.0    285  17  AAW13237                  Human steroidogene
     8      53    100.0    285  17  AAW00814                  Human steroidogene
     9      53    100.0    285  17  AAW132[illegible]         Human steroid acu
    10      53    100.0    285  17  AAW13234                  Human steroidogene
    11      53    100.0    285  17  AAW 323                   Human steroidogene
    12      53    100.0    285  22  AAY72833                  Human steroidogeni
    13      49     92.5    285  17  AAW132[illegible]         Human steroidogene
    14    39.5     74.5    306  21  AAB13[illegible]          Rat tartrate resis
    15    39.5     74.5    332  21  AAB58814                  Lung cancer associ
    16      37     69.8    160  22  AAB76877                  Human lung tumour
    17      36     67.9    122  21  AAG42[illegible]          Arabidopsis thalia
    18      36     67.9    137  21  AAG42280                  Arabidopsis thalia
    19      35     66.0   1072  22  AAB72[illegible]          Human ADAMTS-10 al
    20      35     66.0   1081  22  AAB72[illegible]          Human ADAMTS-10 am
    21      34     64.2     72  21  AAB54[illegible]          Human pancreatic c
    22      34     64.2    917  13  AAR26[illegible]34        gp130. Mus muscul
    23      34     64.2    917  21  AAY55[illegible]          Mouse gp130 protei
    24      34     64.2   4393  22  AAB31089                  Amino acid sequenc
    25      33     62.3      9  15  AAR67[illegible]          N-terminal sequenc
    26      33     62.3      9  20  AAY23[illegible]          N-terminal sequenc
    27      33     62.3     63  21  AAB09[illegible]4         Hepatitis GB virus
    28      33     62.3     99  18  AAW10[illegible]          Protein encoded by
    29      33     62.3    232  20  AAY23[illegible]          Canine hookworm ne
    30      33     62.3    234  14  AAR52[illegible]          Canine hookworm Ne
    31      33     62.3    234  21  AAG36[illegible]          Arabidopsis thalia
    32      33     62.3    241  20  AAY23610                  Canine hookworm ne
    33      33     62.3    248  20  AAY23[illegible]          Canine hookworm ne
    34      33     62.3    450  21  AAG36602                  Arabidopsis thalia
    35      33     62.3    466  21  AAG36[illegible]          Arabidopsis thalia
    36      33     62.3    497  22  AAB74476                  Human protease inh
    37      33     62.3   3011  16  AAR95[illegible]          Hepatitis GB virus
    38      32     60.4      9  15  AAR67766                  Hybrid N-terminal
    39      32     60.4      9  20  AAY235[illegible]         N-terminal sequenc
    40      32     60.4     34  22  AAB66025                  Human secreted pro
    41      32     60.4     83  21  AAB084[illegible]4        Amino acid sequenc
    42      32     60.4    207  21  AAY94954                  Human secreted pro
    43      32     60.4    217  12  AAP[illegible]            [illegible] lipase.
    44      32     60.4    219  21  AAG[illegible]            Arabidopsis thalia
    45      32     60.4    217  21  AAG[illegible]            Arabidopsis thalia


                                 ALIGNMENTS


RESULT   1
AAY72836
ID   AAY72836 standard; peptide; 9 AA.
XX
AC   AAY72836;
XX
DT   21-MAY-2001  (first entry)
XX
DE   Steroidogenic acute regulatory (StAR) peptide #2.
XX
KW   Steroidogenic acute regulatory protein; StAR;
KW   gene therapy; steroid hormone-dependent disorder; LCAH;
KW   lipoid congenital adrenal hyperplasia; sexual maturation;
KW   infertility; androgen-responsive tumour; precocious puberty;
KW   hypogonadotropic hypogonadism; McCune-Albright syndrome;
KW   adrenal-hypoplasia congenita; cytostatic.
XX
OS   Unidentified.
XX
PN   US6194555-B1.
XX
PD   27-FEB-2001.
XX
PF   07-JUN-1996;   96US-0659254.
XX
PR   04-NOV-1995;   95US-0538960.
XX
PA   (CYTE ) UNIV TEXAS TECH HEALTH SCI CENT.
XX
PI   Stocco DM,  Clark BJ;
XX
DR   WPI; 2001-217935/22.
XX
PT   Novel nucleic acid encoding steroidogenic acute regulatory protein,
PT   useful for treating steroid hormone dependent disorders
XX
PS   Disclosure; Column [illegible]; 41pp; English.
XX
CC   The invention relates to steroidogenic acute regulatory proteins
CC   (STAR) and nucleic acid molecules encoding them. STAR proteins
CC   are used for regulating steroidogenesis. More particularly, these
CC   proteins are involved in regulating cholesterol transport into
CC   the mitochondria of a cell. The nucleic acid molecules of the
CC   invention are useful in gene therapy for the treatment of steroid
CC   hormone dependent disorders, particularly lipoid congenital
CC   adrenal hyperplasia (LCAH), infertility, sexual maturation,
CC   androgen responsive tumours, precocious puberty, McCune Albright
CC   syndrome, adrenal hypoplasia congenita or hypogonadotropic
CC   hypogonadism. They are also useful for recombinant preparation
CC   of the steroidogenic acute regulatory proteins.
CC   The present sequence is steroidogenic acute regulatory (STAR)
CC   peptide. This peptide is useful for [illegible]
CC   screening and identifying antigenic determinants of STAR proteins.
XX
SQ   Sequence   9 AA;

  Query Match   [illegible]
  Best Local Similarity   100.0%;  Pred. No. [illegible]
  Matches   [illegible];  Conservative   [illegible];  Mismatches   [illegible]

Qy        1 AFHGPTCMV [illegible]
            ||||||||
Db        1 aehgptcmv [illegible]


RESULT 2
AAW74154
ID   AAW74154 standard; Protein; 14 AA.
XX
AC   AAW74154;
XX
DT   05-MAY-1999  (first entry)
XX
DE   Mouse steroidogenic acute regulatory protein tryptic peptide [illegible]
XX
KW   StAR; steroidogenic acute regulatory protein; McCune Albright [illegible]
KW   steroid hormone dependent disorder; adrenal hypoplasia congenita; LCAH;
KW   lipoid congenital adrenal hyperplasia; infertility; sexual maturation;
KW   androgen responsive tumour; hypogonadism; therapy; [illegible]
XX
OS   Mus sp.
XX
PN   US5872[illegible]-A.
XX
PD   16-FEB-1999.
XX
PF   04-NOV-1995;   [illegible]
XX
PR   04-NOV-1995;   [illegible]
XX
PA   (UYTE ) UNIV TEXAS TECH HEALTH SCI CENT.
XX
PI   Clark BJ,  Stocco DM;
XX
DR   WPI; [illegible]
XX
PT   New nucleic acid encoding steroidogenic acute regulatory protein -
PT   useful for treatment of steroid hormone dependent disorders, e.g.
PT   lipoid congenital adrenal hyperplasia, infertility and
PT   androgen responsive tumours
XX
PS   Example 2; Column 26; 41pp; English.
XX
CC   This sequence is a tryptic peptide fragment of the steroidogenic acute
CC   regulatory protein (STAR) of the invention. The STAR protein sequence and
CC   its products or derivatives may be used in [illegible]
CC   [illegible] preparation of STAR and in the treatment of steroid hormone dependent
CC   disorders such as lipoid congenital adrenal hyperplasia (LCAH)
CC   infertility, sexual maturation, androgen responsive tumours, precocious
CC   puberty, McCune Albright syndrome, adrenal hypoplasia congenita and
CC   hypogonadism [illegible] STAR [illegible]
CC   progesterone levels and [illegible] risk of spontaneous abortion.
XX
SQ   Sequence   14 AA;

  Query Match   [illegible]   Length 14;
  Best Local Similarity   [illegible];  Pred. No. [illegible]
  Matches   [illegible];  Conservative   [illegible];  Mismatches   [illegible]

Qy        1 AFH[illegible]CMV [illegible]

Db        1 [illegible]


RESULT [illegible]
AAY[illegible]
ID   AAY[illegible] standard; peptide; [illegible] AA.
XX
AC   AAY[illegible];
XX
DT   [illegible]-MAY-[illegible]  (first entry)
XX
DE   Mouse tryptic peptide #[illegible] for analysis of [illegible] StAR protein.
XX
KW   Mouse steroidogenic acute regulatory protein; StAR;
KW   gene therapy; steroid hormone dependent disorder; LCAH;
KW   lipoid congenital adrenal hyperplasia; sexual maturation;
KW   infertility; androgen responsive tumour; precocious puberty;
KW   hypogonadotropic hypogonadism; McCune Albright syndrome;
KW   adrenal hypoplasia congenita; cytostatic; tryptic peptide.
XX
OS   Mus sp.
XX
PN   [illegible]
XX
PD   [illegible]
XX
PF   [illegible] NOV [illegible]
XX
PR   [illegible] NOV [illegible]
XX
PA   (UYTE ) UNIV TEXAS TECH HEALTH SCI CENT.
XX
PI   [illegible] M[illegible] Clark [illegible]
XX
DR   WPI; [illegible]
XX
PT   New nucleic acid encoding steroidogenic acute regulatory protein -
PT   useful for treating steroid hormone dependent disorders
XX
PS   Example [illegible]; [illegible]; 41pp; English.
XX
CC   The present invention relates to steroidogenic acute regulatory proteins
CC   (STAR) and nucleic acid molecules encoding them. STAR proteins
CC   are used for regulating steroidogenesis. More particularly these
CC   proteins are involved in regulating cholesterol transport into
CC   the mitochondria of a cell. The nucleic acid molecules of the
CC   invention are useful in gene therapy for the treatment of steroid
CC   hormone dependent disorders, particularly lipoid congenital
CC   adrenal hyperplasia (LCAH), infertility, sexual maturation,
CC   androgen responsive tumours, precocious puberty, McCune Albright
CC   syndrome, adrenal hypoplasia congenita or hypogonadotropic
CC   hypogonadism. They are also useful for recombinant preparation
CC   of the steroidogenic acute regulatory proteins.
CC   The present sequence is tryptic peptide #[illegible]. This peptide is
CC   used in the cloning and sequence analysis of cDNA encoding the
CC   [illegible] STAR protein. This peptide is also used in the
CC   analysis of the STAR protein sequence. Peptide #[illegible] is useful
CC   for designing oligonucleotides for screening and identifying
CC   antigenic determinants of StAR protein.
XX
SQ   Sequence   14 AA;

  Query Match              100.0%;  Score 53;  DB 22;  Length 14;
  Best Local Similarity    100.0%;  Pred. No. 0.0045;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         1 AEHGPTCMV 9
             |||||||||
Db         1 aehgptcmv 9


RESULT   4
AAW13238
ID   AAW13238 standard; Protein; 284 AA.
XX
AC   AAW13238;
XX
DT   21-MAY-1997  (first entry)
XX
[illegible]   Human steroidogenesis acute regulatory protein mutant delta Arg272.
XX
KW   Human; steroidogenesis; acute regulatory protein; hStAR; analysis;
KW   mutation; detection; prenatal; genetic defect; congenital; protein;
KW   lipoid adrenal hyperplasia; treatment; prevention; gene;
KW   replacement therapy; hypercholesterolaemia; mutant.
XX
OS   Homo sapiens.
XX
PN   WO9629338-A1.
XX
PD   26-SEP-1996.
XX
PF   22-MAR-1996;   96WO-US03896.
XX
PR   23-MAR-1995;   95US-0410540.
XX
PA   (REGC ) UNIV CALIFORNIA.
PA   (UYPE-) UNIV PENNSYLVANIA.
XX
PI   Lin D,  Miller WL,  Strauss JF;
XX
DR   WPI; 1996 443130/44.
XX
PT   Isolated human steroidogenesis acute regulatory protein gene - used
PT   for detection of mutation(s) of this gene that cause congenital
PT   lipoid adrenal hyperplasia
XX
[illegible]   Claim 16; Page -; 89pp; English.
XX
CC   The present sequence is the human steroidogenesis acute regulatory
CC   protein (hStAR), congenital lipoid adrenal hyperplasia (CAH)
CC   associated mutant, delta Arg272 (i.e. wild type residue Arg272 has
CC   been deleted). Analysing the hStAR gene for such a mutation is
CC   useful in the detection (e.g. prenatally) of CAH associated
CC   genetic defects, or its transmission to children. CAH can be
CC   treated by protein or gene replacement therapy, which can also be
CC   used to prevent or treat hypercholesterolaemia.
XX
SQ   Sequence   284 AA;

  Query Match              100.0%;  Score 53;  DB 17;  Length 284;
  Best Local Similarity    100.0%;  Pred. No. 0.675;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         1 AEHGPTCMV 9
             |||||||||
Db       218 aehgptcmv 226
```

```
RESULT   5
AAW74153
ID   AAW74153 standard; Protein; 284 AA.
XX
AC   AAW74153;
XX
DT   05 MAY 1999  (first entry)
XX
DE   Mouse steroidogenic acute regulatory protein.
XX
KW   StAR; steroidogenic acute regulatory protein; McCune-Albright syndrome;
KW   steroid hormone dependent disorder; adrenal hypoplasia congenita; LCAH;
KW   lipoid congenital adrenal hyperplasia; infertility; sexual maturation;
KW   androgen-responsive tumours; hypogonadism; therapy; progesterone level.
XX
OS   Mus sp.
XX
PN   US5872230-A.
XX
PD   16-FEB-1999.
XX
PF   04-NOV-1995;   95US-0538960.
XX
PR   04-NOV-1995;   95US-0538960.
XX
PA   (UYTE-) UNIV TEXAS TECH HEALTH SCI CENT.
XX
PI   Clark BJ,  Stocco DM;
XX
DR   WPI; 1999 166720/14.
DR   N-PSDB; AAX18169, AAX18170.
XX
PT   New nucleic acid encoding steroidogenic acute regulatory protein -
PT   useful for treatment of steroid hormone-dependent disorders such as
PT   lipoid congenital adrenal hyperplasia, infertility and
PT   androgen responsive tumours
XX
PS   Claim 1; Column 37-38; 25pp; English.
XX
CC   This sequence is the steroidogenic acute regulatory protein (StAR)
CC   of the invention. The StAR coding sequence and its products or
CC   derivatives may be used as probes and primers for PCR hybridisation
CC   studies and in the treatment of steroid hormone dependent disorders such
CC   as lipoid congenital adrenal hyperplasia (LCAH), infertility, sexual
CC   maturation, androgen-responsive tumours, precocious puberty,
CC   McCune-Albright syndrome, adrenal hypoplasia congenita, and hypogonadism.
CC   Monitoring levels of StAR may also indicate reduced progesterone levels
CC   and so a risk of spontaneous abortion.
XX
SQ   Sequence   284 AA;

  Query Match              100.0%;  Score 53;  DB 20;  Length 284;
  Best Local Similarity    100.0%;  Pred. No. 0.675;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         1 AEHGPTCMV 9
             |||||||||
Db       217 aehgptcmv 225


RESULT   6
AAY72831
ID   AAY72831 standard; Protein; 284 AA
XX
AC   AAY72831;
XX
DT   31 MAY 2001  (first entry)
XX
DE   Mouse steroidogenic acute regulatory (StAR) protein.
XX
KW   Mouse; steroidogenic acute regulatory protein; StAR;
KW   gene therapy; steroid hormone dependent [illegible] StAR
KW   lipoid congenital adrenal hyperplasia; [illegible]
KW   infertility; androgen responsive tumour; precocious puberty;
KW   hypogonadotropic hypogonadism; McCune Albright syndrome;
KW   adrenal hypoplasia congenita; [illegible]
XX
OS   Mus sp.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..47
FT                   /label= Signal peptide
FT   Cleavage site   [illegible]
FT                   /note= "Site for mitochondrial protease
FT                   matrix processing protease [illegible]
FT                   intermediate processing peptide"
FT   Protein         48..284
FT                   /label= Mature steroidogenic acute regulatory protein
FT   Cleavage site   [illegible]
FT                   /note= "Site for mitochondrial protease
FT                   matrix processing protease and [illegible]
FT                   intermediate processing peptide"
FT   Modified site   [illegible]
FT                   /note= "PKA phosphorylation site"
FT   Modified site   [illegible]
FT                   /note= "[illegible]K phosphorylation site"
FT   Region          [illegible]
FT                   /label= Tryptic peptide #2
FT   Region          [illegible]
FT                   /note= "This sequence is used to raise antibodies
FT                   against StAR protein"
FT   Region          [illegible]..[illegible]
FT                   /label= Tryptic peptide #1
FT   Modified site   195
FT                   /note= "PKA phosphorylation site"
FT   Region          [illegible]
FT                   /label= Tryptic peptide #3
FT   Modified site   [illegible]
FT                   /note= "[illegible] kinase [illegible] phosphorylation site"
FT   Modified site   238
FT                   /note= "PKC phosphorylation site"
XX
PN   US619455[illegible] B1.
XX
PD   27 FEB 2001.
XX
PF   07 JUN 1996;   96US-[illegible]
XX
PR   04 NOV 1995;   95US-[illegible]
XX
PA   (UYTE ) UNIV TEXAS TECH HEALTH SCI CENT.
XX
PI   Stocco DM, Clark BJ;
XX
DR   WPI; 2001-21[illegible]
DR   N-PSDB; AAD0[illegible]   AAD[illegible]
XX
PT   Novel nucleic acid encoding steroidogenic acute regulatory protein,
PT   useful for treating steroid hormone dependent disorders
XX
PS   Example 2; Fig 2; [illegible]pp; English.
XX
CC   The invention relates to steroidogenic acute regulatory protein
CC   (StAR) and nucleic acid molecules encoding them. StAR proteins
CC   are used for regulating steroidogenesis. More particularly, StAR
CC   proteins are involved in regulation of cholesterol transport into
CC   the mitochondria of a cell. The nucleic acid molecules of the
CC   invention are useful in gene therapy for the treatment of steroid
CC   hormone dependent disorders, particularly, lipoid congenital
CC   adrenal hyperplasia (LCAH), infertility, sexual [illegible],
CC   androgen responsive tumours, precocious puberty, McCune Albright
CC   syndrome, adrenal hypoplasia congenita, [illegible]
CC   hypogonadism. They are also useful for recombinant production
CC   of the steroidogenic acute regulatory protein.
```

```
CC   The present sequence is mouse steroidogenic acute regulatory
CC   protein (StAR).
XX
SQ   Sequence   284 AA;

  Query Match            [illegible];  Score [illegible];  DB [illegible];  Length 284;
  Best Local Similarity  [illegible];  Pred. No. [illegible];
  Matches [illegible];  Conservative [illegible];  Mismatches [illegible];  Indels [illegible];  Gaps [illegible];

Qy        [illegible]

Db        [illegible]


RESULT [illegible]
AAW[illegible]
ID   AAW[illegible] standard; protein; 28[illegible] AA.
XX
AC   AAW[illegible];
XX
DT   [illegible] MAY [illegible] (first entry)
XX
DE   Human steroidogenic acute regulatory protein mutant Leu2[illegible]Pro.
XX
KW   Human steroidogenic acute regulatory protein; hStAR; analysis;
KW   mutation; detection; prenatal; genetic defect; congenital; protein;
KW   lipoid adrenal hyperplasia; treatment; prevention; gene
KW   replacement therapy; steroids; esters; hormone; mutant.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference [illegible]
FT                   /note= "wild type Leu replaced with Pro"
XX
PN   WO9[illegible]-A1.
XX
PD   [illegible] OCT [illegible].
XX
PF   [illegible] MAR [illegible];   96WO-US[illegible]
XX
PR   [illegible] MAR [illegible];   95US-[illegible]
XX
PA   (REGC ) UNIV CALIFORNIA.
PA   (UYPE ) UNIV PENNSYLVANIA.
XX
PI   Lin D, Miller WL, Strauss JF;
XX
DR   WPI; 199[illegible]-[illegible]/43.
DR   N-PSDB; AA[illegible]
XX
PT   Isolated human steroidogenic acute regulatory protein gene - used
PT   for detection of mutation(s) in this gene that cause congenital
PT   lipoid adrenal hyperplasia
XX
PS   Claim 15; Page [illegible]; [illegible]pp; English.
XX
CC   The present sequence is the human steroidogenic acute regulatory
CC   protein (hStAR) congenital lipoid adrenal hyperplasia (CAH)
CC   associated mutant [illegible]. Analysis of the hStAR gene for such a
CC   mutation is useful in the detection (e.g. prenatally) of CAH
CC   [illegible] to detect [illegible] its transmission to children. CAH
CC   can be treated by protein or gene replacement therapy, which can
CC   also be used to prevent or treat hypercholesterolaemia.
XX
SQ   Sequence   28[illegible] AA;

  Query Match            [illegible];  Score [illegible];  DB [illegible];  Length 28[illegible];
  Best Local Similarity  [illegible];  Pred. No. [illegible];
  Matches [illegible];  Conservative [illegible];  Mismatches [illegible];  Indels [illegible];  Gaps [illegible];
```

```
Qy       1 AEHGPTCMV 9
           |||||||||
Db     218 aehqptcmv 226


RESULT    8
AAW00814
ID   AAW00814 standard; Protein; 285 AA.
XX
AC   AAW00814;
XX
DT   21-MAY-1997  (first entry)
XX
DE   Human steroidogenesis acute regulator, protein.
XX
KW   Human; steroidogenesis; acute regulator; protein; hStAR; analysis;
KW   mutation; detection; prenatal; genetic defect; congenital; protein;
KW   lipoid adrenal hyperplasia; treatment; prevention; gene;
KW   replacement therapy; hypercholesterolaemia.
XX
OS   Homo sapiens.

     Key             Location/Qualifiers
     Region          5
FT                   /note= "potential site for protein kinase C
FT                          mediated phosphorylation"
FT   Region          13
FT                   /note  "potential site for protein kinase C
FT                          mediated phosphorylation"
FT   Region          56..57
FT                   /note= "potential site for protein kinase A
FT                          mediated phosphorylation"
FT   Region          100
FT                   /note= "potential site for protein kinase A
FT                          mediated phosphorylation"
FT   Region          110
FT                   /note  "potential site for protein kinase A
FT                          mediated phosphorylation"
FT   Region          186
FT                   /note= "potential site for protein kinase C
FT                          mediated phosphorylation"
FT   Region          195..196
FT                   /note= "potential site for protein kinase A
FT                          mediated phosphorylation"
FT   Region          277
FT                   /note= "potential site for protein kinase A
FT                          mediated phosphorylation"
XX
PN   WO9629338-A1.

     26-SEP-1996.

PF   22-MAR-1996;   96WO-US03896.
XX
PR   23-MAR-1995;   95US-0410540.
XX
PA   (REGC ) UNIV CALIFORNIA.
PA   (UYPE-) UNIV PENNSYLVANIA.
XX
PI   Lin D,  Miller WL,  Strauss JF;
XX
DR   WPI; 1996-443130/44.
DR   N-PSDB; AAT33650.
XX
PT   Isolated human steroidogenesis acute regulatory protein gene - used
PT   for detection of mutation(s) of this gene that cause congenital
PT   lipoid adrenal hyperplasia
XX
PS   Claim 1; Pages 62-63; 89pp; English.
XX
CC   The present sequence is the human steroidogenesis acute
CC   regulatory protein (hStAR). The hStAR gene can be analysed for
CC   mutations to detect (e.g. prenatally) genetic defects associated
CC   with congenital lipoid adrenal hyperplasia (CAH), or its
CC   transmission to children. CAH can be treated by protein or gene
CC   replacement therapy, which can also be used to prevent or treat
CC   hypercholesterolaemia.
CC   A human adrenal cortex cDNA library was screened with a mouse StAR
CC   probe to isolate a 1.6 kb insert, including an ORF for a 285
CC   residue protein. When it was cloned into pSPORT and expressed in
CC   COS-1 cells cotransfected with pP450scc and pADX, it increased the
CC   level of pregnenolone synthesis from cholesterol or
CC   20 alpha-hydroxycholesterol.
XX
SQ   Sequence   285 AA;

  Query Match             100.0%;  Score 53;  DB 17;  Length 285;
  Best Local Similarity   100.0%;  Pred. No. 0.075
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AEHGPTCMV 9
           |||||||
Db     218 aehqptcmv 226


RESULT    9
AAW13233
ID   AAW13233 standard; Protein; 285 AA.
XX
AC   AAW13233;
XX
DT   21 MAY 1997  (first entry)
XX
DE   Human steroidogenesis acute regulatory protein mutant Glu169Gly.
XX
KW   Human, steroidogenesis, acute regulatory protein, hStAR; analysis;
KW   mutation; detection; prenatal; genetic defect; congenital; protein;
KW   lipoid adrenal hyperplasia; treatment; prevention; gene;
KW   replacement therapy; hypercholesterolaemia; mutant.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 169
FT                   /note  "wild type Glu substituted with Gly"
XX
PN   WO9629338 A1.
XX
PD   26-SEP-1996.
XX
PF   22-MAR-1996;   96WO-US03896.
XX
PR   23-MAR-1995;   95US-0410540.
XX
PA   (REGC ) UNIV CALIFORNIA.
PA   (UYPE-) UNIV PENNSYLVANIA.
XX
PI   Lin D,  Miller WL,  Strauss JF;
XX
DR   WPI; 1996-443130/44.
DR   N-PSDB; AAT62018.
XX
PT   Isolated human steroidogenesis acute regulatory protein gene - used
PT   for detection of mutation(s) of this gene that cause congenital
PT   lipoid adrenal hyperplasia
XX
PS   Claim 13; Page  ; 89pp; English.
XX
CC   The present sequence is the human steroidogenesis acute regulatory
CC   protein (hStAR), congenital lipoid adrenal hyperplasia (CAH)
CC   associated mutant, Glu169Gly. Analysing the hStAR gene for such a
CC   mutation is useful in the detection (e.g. prenatally) of CAH
CC   associated genetic defects, or its transmission to children. CAH
CC   can be treated by protein or gene replacement therapy, which can
CC   also be used to prevent or treat hypercholesterolaemia.
```

```
XX
SQ   Sequence   [illegible] AA;


  Query Match             [illegible]
  Best Local Similarity   [illegible]
  Matches    9; [illegible] Mismatches [illegible]

Qy          1 AEHGPICMV [illegible]
                [illegible]
Db        218 [illegible]


RESULT 10
AAW13234
ID   AAW13234 standard; Protein; 28[illegible] AA.
XX
     AAW13234;
XX
DT   21 MAY 1997  (first entry)
XX
DE   Human steroidogenesis acute regulatory protein [illegible]
XX
KW   Human; steroidogenesis acute regulatory protein; [illegible]
KW   mutation; detection; prenatal; genetic defect; [illegible]
KW   lipoid adrenal hyperplasia; treatment; prevention; [illegible]
KW   replacement therapy; hypercholesterolaemia; mutant
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc difference [illegible]
FT                   /note= "wild type Arg substituted [illegible]"
XX
PN   WO9629[illegible] A1.
XX
PD   26 SEP 1996.
XX
PF   22 MAR 1996;  [illegible]
XX
PR   23 MAR 1995;  [illegible]
XX
PA   (REGC ) UNIV CALIFORNIA.
PA   (UYPE ) UNIV PENNSYLVANIA.
XX
     Lin D,  Miller W,  [illegible]
XX
     WPI; 1996-443[illegible]
DR   N-PSDB; AAT6[illegible]
XX
PT   Isolated human steroidogenesis acute regulatory protein [illegible]
PT   for detection [illegible]
PT   lipoid adrenal hyperplasia
XX
PS   Claim 13; Page [illegible]pp; [illegible]
XX
CC   The present sequence [illegible] the human ster[illegible]
CC   protein (hStAR), congenital lipoid adrenal hyperplasia [illegible]
CC   associated mutant Ar[illegible]. Analysing the hStAR gene [illegible]
CC   mutation is useful in the detection (e[illegible]
CC   associated genetic defects [illegible]
CC   can be treated by protein or gene replacement therapy [illegible]
CC   also be used to prevent or treat hypercholesterolaemia.
XX
SQ   Sequence   28[illegible] AA;


  Query Match             [illegible]
  Best Local Similarity   [illegible]
  Matches    9; Conservative [illegible] Mismatches [illegible] Indels [illegible] Gaps [illegible]

Qy          1 AEHGPICMV 9
                [illegible]
```

```
Db   [illegible]


RESULT 11
AAW1[illegible]
ID   AAW1[illegible] standard; Protein; 28[illegible] AA.
XX
AC   AAW1[illegible]
XX
DT   [illegible] MAY [illegible]  (first entry)
XX
DE   Human steroidogenesis acute regulatory protein mutant Glu[illegible]Cys.
XX
KW   Human; steroidogenesis acute regulatory protein; hStAR; analysis;
KW   [illegible] detection; [illegible] genetic defect; congenital; protein;
KW   lipoid adrenal hyperplasia; treatment; prevention; gene;
KW   [illegible]
XX
OS   [illegible]
XX
FH   Key             [illegible]
FT   Misc difference [illegible]
FT                   [illegible] "wild type [illegible] substituted with Cys"
XX
PN   W[illegible] A1.
XX
PD   [illegible]
XX
PF   [illegible] MAR [illegible]
XX
PR   [illegible] MAR [illegible]
XX
PA   (REGC ) UNIV CALIFORNIA.
PA   [illegible] UNIV PENNSYLVANIA.
XX
PI   [illegible] Miller W, [illegible]
XX
DR   WPI; [illegible]
DR   N-PSDB; AAT[illegible]
XX
PT   Isolated human steroidogenesis acute regulatory protein gene [illegible] used
PT   [illegible] congenital
PT   lipoid adrenal hyperplasia
XX
PS   Claim [illegible]; Page [illegible]pp; English.
XX
CC   The present sequence is the human steroidogenesis acute regulatory
CC   protein (hStAR), [illegible] lipoid adrenal hyperplasia (CAH)
CC   associated mutant, [illegible]Cys. Analysing the hStAR gene for such a
CC   mutation [illegible] in the detection (e.g. prenatally) of CAH
CC   associated genetic defects, or its transmission to children. CAH
CC   can be treated by protein or gene replacement therapy, which can
CC   also be used to prevent or treat hypercholesterolaemia.
XX
SQ   Sequence   [illegible] AA;


  Query Match             [illegible] Length 28[illegible];
  Best Local Similarity   [illegible] Pred. No. [illegible]
  Matches [illegible] Conservative [illegible] Mismatches [illegible] Indels [illegible] Gaps [illegible]

Qy          [illegible] MV [illegible]
                [illegible]
Db        [illegible]


RESULT 12
AAY[illegible]
ID   AAY[illegible] standard; [illegible] AA
XX
AC   AAY[illegible]
XX
DT   [illegible] MAY [illegible]  (first entry)
```

```
XX
DE   Human steroidogenic acute regulatory (StAR) protein.
XX
KW   Human; steroidogenic acute regulatory protein; StAR;
KW   gene therapy; steroid hormone-dependent disorder; LCAH;
KW   lipoid congenital adrenal hyperplasia; sexual maturation;
KW   infertility; androgen responsive tumour; precocious puberty;
KW   hypogonadotropic hypogonadism; McCune-Albright syndrome;
KW   adrenal-hypoplasia congenita; cytostatic.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 156
FT                   /note  "Encoded by CAC; this feature corresponds
FT                   to the translation exception for the sequence shown
FT                   in (AAD02909)"
FT   Modified-site   59..60
FT                   /note  "PKA phosphorylation site"
FT   Modified-site   68..69
FT                   /note- "CK phosphorylation site"
FT   Modified-site   195
FT                   /note  "PKA phosphorylation site"
FT   Modified-site   233..234
FT                   /note- "P34CDC2 kinase (?D) phosphorylation site"
FT   Modified-site   238
FT                   /note  "PKC phosphorylation site"
FT   Modified-site   277
FT                   /note- "PKA phosphorylation site"
XX
PN   US6194555-B1.
XX
PD   27-FEB-2001.
XX
PF   07-JUN-1996;   96US-0659254.
XX
PR   04-NOV-1995;   95US-0538960.
XX
PA   (UYTE-) UNIV TEXAS TECH HEALTH SCI CENT.
XX
PI   Stocco DM,  Clark BJ;
XX
DR   WPI; 2001-217936/22.
DR   N-PSDB; AAD02909, AAD02912.
XX
PT   Novel nucleic acid encoding steroidogenic acute regulatory protein,
PT   useful for treating steroid hormone-dependent disorders -
XX
PS   Example 14; Fig 5; 41pp; English.
XX
CC   The invention relates to steroidogenic acute regulatory protein
CC   (StAR) and nucleic acid molecules encoding them. StAR proteins
CC   are used for regulating steroidogenesis. More particularly these
CC   proteins are involved in regulating cholesterol transport into
CC   the mitochondria of a cell. The nucleic acid molecules of the
CC   invention are useful in gene therapy for the treatment of steroid
CC   hormone-dependent disorders, particularly lipoid congenital
CC   adrenal hyperplasia (LCAH), infertility, sexual maturation,
CC   androgen-responsive tumours, precocious puberty, McCune-Albright
CC   syndrome, adrenal-hypoplasia congenita or hypogonadotropic
CC   hypogonadism. They are also useful for recombinant preparation
CC   of the steroidogenic acute regulatory protein.
CC   The present sequence is human steroidogenic acute regulatory
CC   protein (StAR).
XX
SQ   Sequence   285 AA;

  Query Match             100.0%;  Score 33;  DB 22;  Length 285;
  Best Local Similarity   100.0%;  Pred. No. 0.075;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 AEHGPTCMV 9
              |||||||||
Db      218 aehgptcmv 226


RESULT  13
AAW13236
ID   AAW13236 standard; Protein; 285 AA
XX
AC   AAW13236;
XX
DT   21-MAY 1997  (first entry)
XX
DE   Human steroidogenesis acute regulatory protein mutant Ala218Val.
XX
KW   Human, steroidogenesis, acute regulatory protein, hStAR; analysis;
KW   mutation; detection; prenatal; genetic defect; congenital; protein;
KW   lipoid adrenal hyperplasia; treatment; prevention; gene;
KW   replacement therapy; hypercholesterolaemia, mutant.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 218
FT                   /note  "wild type Ala substituted with Val"
XX
PN   WO9629338-A1.
XX
PD   26-SEP-1996.
XX
PF   22-MAR-1996;   96WO-US03896.
XX
PR   23-MAR-1995;   95US-0410540.
XX
PA   (REGC ) UNIV CALIFORNIA.
PA   (UYPE-) UNIV PENNSYLVANIA.
XX
PI   Lin D,  Miller WL,  Strauss JF;
XX
DR   WPI; 1996-443140/44.
DR   N-PSDB; AAT62021.
XX
PT   Isolated human steroidogenesis acute regulatory protein gene - used
PT   for detection of mutation(s) of this gene that cause congenital
PT   lipoid adrenal hyperplasia
XX
PS   Claim 13; Page  ; 89pp; English.
XX
CC   The present sequence is the human steroidogenesis acute regulatory
CC   protein (hStAR), congenital lipoid adrenal hyperplasia (CAH)
CC   associated mutant, Ala218Val. Analysing the hStAR gene for such a
CC   mutation is useful in the detection (e.g. prenatally) of CAH
CC   associated genetic defects, or its transmission to children. CAH
CC   can be treated by protein or gene replacement therapy, which can
CC   also be used to prevent or treat hypercholesterolaemia.
XX
SQ   Sequence   285 AA;

  Query Match             92.5%;  Score [illegible];  DB 17;  Length 285;
  Best Local Similarity   100.0%;  Pred. No. 0.37
  Matches    8;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        2 EHGPTCMV 9
              ||||||||
Db      219 ehgptcmv 226


RESULT  14
AAB13356
ID   AAB13356 standard; protein; 30? AA.
XX
AC   AAB13356;
XX
```

```
DT   21 NOV 2000  (first entry)
XX
DE   Rat tartrate resistant purple acid phosphatase.
XX
KW   Rat; tartrate resistant purple acid phosphatase; [illegible]
KW   purple acid phosphatase; PAP; [illegible]
KW   bone metabolism disorder; osteoporosis; [illegible]
KW   protein co-ordinate data.
XX
OS   Rattus sp.
XX
PN   WO200037620-A1.
XX
PD   29-JUN-2000.
XX
PF   25-NOV-1999;   99WO-[illegible]
XX
     18-DEC-1998;   98SE-004438.
XX
PA   (PHAA ) PHARMACIA & UPJOHN AB.
XX
PI   Uppenberg J;
XX
DR   WPI; 2000-442557/38.
XX
PT   Crystalline form of tartrate resistant acid purple [illegible]
PT   activated by cleavage - useful for treatment of diseases or degenerative
PT   conditions resulting in increased bone resorption [illegible]
XX
PS   Disclosure; [illegible]  74pp; English.
XX
CC   The present sequence is rat tartrate resistant [illegible]
CC   (TRAP). TRAP belongs to a group of enzymes called the [illegible]
CC   phosphatases (PAPs), which are characterised [illegible]
CC   at the active site. TRAP is expressed predominantly [illegible]
CC   cells (osteoclasts), where it may mediate dephosphorylation [illegible]
CC   matrix proteins such as osteopontin and bone sialoprotein. A crystalline
CC   form of activated TRAP was prepared by [illegible]
CC   protease and then crystallising the product. [illegible]
CC   used for structure based drug design for specific [illegible]
CC   and inhibitors of TRAP activity. These are useful [illegible]
CC   of diseases or degenerative conditions resulting in increased bone
CC   resorption, such as tissue damage, [illegible]
CC   osteoporosis.
XX
SQ   Sequence   304 AA;

  Query Match             [illegible]
  Best Local Similarity   [illegible];  Pred. No. [illegible]
  Matches    8;  Conservative  [illegible];  Mismatches  [illegible]

Qy          1 AEHGPI MV [illegible]
              |||||| |:
Db        193 aehaptr[illegible]


RESULT 15
AAB58146
ID   AAB58146 standard; Protein; [illegible] AA.
XX
AC   AAB58146;
XX
DT    4 MAR 2001  (first entry)
XX
DE   Lung cancer associated polypeptide sequence [illegible]
XX
KW   Human; lung cancer associated protein; [illegible]
KW   cardioactive; immunomodulatory; muscular; [illegible]
KW   gastrointestinal; nephrotropic; antiinfective; [illegible]
KW   antibacterial; diagnosis; neural disorder; [illegible]
KW   proliferative disorder; wound healing; [illegible]
XX
OS   Homo sapiens.
XX
PN   WO[illegible]-A1.
XX
PD   [illegible]
XX
PF   [illegible] MAR [illegible]  [illegible]W[illegible]
XX
PR   [illegible] MAR [illegible]
XX
PA   (HUMA-) HUMAN GENOME SCIENCES INC.
PA   (ROSE/) ROSEN C A.
XX
PI   Ruben SM;
XX
DR   WPI; [illegible]
DR   N-PSDB; AAF[illegible]
XX
PT   [illegible] associated [illegible] referred to as lung cancer
PT   antigens - useful for treatment, prevention and diagnosis of disorders
PT   such as lung cancer
XX
PS   [illegible] English.
XX
CC   Polypeptide sequences AAB[illegible] AAB[illegible] encode human lung cancer
CC   associated proteins represented in AAB[illegible] AAB58148. Lung cancer
CC   associated proteins and polynucleotide sequences, their agonists and
CC   antagonists may have [illegible] protective [illegible]
CC   immunomodulatory, muscular, [illegible] gastrointestinal and
CC   nephrotropic [illegible] antiinfective [illegible] antibacterial
CC   activity. The invention also includes antibodies specific for the
CC   proteins or polynucleotide sequences. The lung cancer associated
CC   polynucleotide sequences may be used for detection of lung cancer,
CC   chromosome identification, chromosome markers, and for numerous other
CC   diagnostic or research purposes. The proteins may be used to treat
CC   disorders such as neural, immune, muscular, reproductive,
CC   gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
CC   disorders. The proteins may also be used in the treatment of wounds and
CC   infectious diseases. Polynucleotide sequences AAF[illegible] AAF[illegible] and
CC   peptide AAB58146 are used in the course of the invention for the
CC   identification and characterisation of the polynucleotide and protein
CC   sequences.
XX
SQ   Sequence   [illegible] AA;

  Query Match             [illegible];  Length [illegible];
  Best Local Similarity   [illegible];  Pred. No. [illegible]
  Matches    [illegible];  Conservative  [illegible];  Mismatches  [illegible];  Indels  [illegible];  Gaps  [illegible]

Qy          1 AEHGPI MV [illegible]

Db        [illegible] aehap[illegible]


Search completed: August 22, 2001, [illegible]
[illegible]
```

```
                    GenCore version 4.5
           Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on:         August 22, 2001, 14:30:12 ; Search time 55.29 Seconds
                                           (without alignments)
                                           3.352 Million cell updates/sec

Title:          US-08-659-254-7
Perfect score:  53
Sequence:       1 AEHGPTCMV 9

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       197339 seqs, 20590346 residues

Total number of hits satisfying chosen parameters:      197339

  mum DB seq length: 0
  mum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       Issued_Patents_AA:*
               1:  /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
               2:  /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
               3:  /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
               4:  /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
               5:  /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
               6:  /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                              SUMMARIES
                 %
Result          Query
  No.   Score   Match Length DB  ID                      Description
-----------------------------------------------------------------------------
    1      53   100.0      9  2  US-08-538-960-7         Sequence 7, Appl
    2      53   100.0      9  4  US-08-659-254-7         Sequence 7, Appl
    3      53   100.0     14  2  US-08-538-960-3         Sequence 3, Appl
    4      53   100.0     14  4  US-08-659-254-3         Sequence 3, Appl
    5      53   100.0    276  4  US-08-659-254-2         Sequence 2, Appl
    6      53   100.0    284  2  US-08-538-960-2         Sequence 2, Appl
    7      53   100.0    285  1  US-08-4[illegible]-2    Sequence 2, Appl
    8      53   100.0    285  4  US-08-659-254-18        Sequence 18, Appl
    9      35    66.0   2556  1  US-09-1[illegible]-432-17   Sequence 17, Appl
   10      35    66.0   2556  1  US-08-[illegible]-190A-20  Sequence 20, Appl
   11      35    66.0   2556  3  US-08-5[illegible]-484-20  Sequence 20, Appl
   12      33    62.3      9  1  US-08-17[illegible]-510B-40  Sequence 40, Appl
   13      33    62.3      9  1  US-08-458-218-40        Sequence 40, Appl
   14      33    62.3      9  2  US-08-4[illegible]-497-40   Sequence 40, Appl
   15      33    62.3     99  2  US-08-858-767-28        Sequence 28, Appl
   16      33    62.3     99  2  US-08-8[illegible]-128-28   Sequence 28, Appl
   17      33    62.3    232  1  US-08-17[illegible]-510B-84  Sequence 84, Appl
   18      33    62.3    232  1  US-08-4[illegible]-218-82   Sequence 82, Appl
   19      33    62.3    232  2  US-08-4[illegible]-497-81   Sequence 81, Appl
   20      32    60.4      9  1  US-08-173-510B-39       Sequence 39, Appl
   21      32    60.4      9  1  US-08-4[illegible]-218-39   Sequence 39, Appl
   22      32    60.4      9  2  US-08-4[illegible]-497-39   Sequence 39, Appl
   23      32    60.4    213  1  US-07-[illegible]-678-2    Sequence 2, Appl
   24      32    60.4    434  3  US-08-414-644-23        Sequence 23, Appl
   25      32    60.4    434  3  US-08-414-644-31        Sequence 31, Appl
   26      32    60.4    434  3  US-08-414-644-40        Sequence 40, Appl
   27      32    60.4    434  3  US-08-414-644-46        Sequence 46, Appl
   28      32    60.4    484  3  US-08-3[illegible]-322C-21  Sequence 21, Appl
   29      32    60.4    484  3  US-08-582-776C-36       Sequence 36, Appl
   30      32    60.4    484  3  US-08-4[illegible]-321B-[illegible]  Sequence 3[illegible], Appl
   31      32    60.4    485  3  US-09-[illegible]-378-10   Sequence 10, Appl
   32      32    60.4    531  2  US-08-[illegible]-114A-4   Sequence 4, Appli
   33      32    60.4    531  4  US-08-84[illegible]-281A-4  Sequence 4, Appli
   34      32    60.4    536  2  US-08-97[illegible]-114A-5  Sequence 5, Appli
   35      32    60.4    536  4  US-08-84[illegible]-281A-5  Sequence 5, Appli
   36      32    60.4   1251  5  PCT-US95-02251-3        Sequence 3, Appli
   37      32    60.4   1252  1  US-08-[illegible]-78[illegible]  Sequence [illegible], Appli
   38      32    60.4   1252  2  US-08-[illegible]-65[illegible]-3  Sequence 3, Appli
   39      32    60.4   1253  3  US-08-37[illegible]-722B-4  Sequence 4, Appli
   40      32    60.4   1404  2  US-08-[illegible]-15[illegible]-2  Sequence 2, Appli
   41      32    60.4   1404  3  US-08-[illegible]-72[illegible]A-2  Sequence 2, Appli
   42      32    60.4   2544  2  US-08-[illegible]-626A-32  Sequence 32, Appl
   43      31    58.5    357  2  US-08-47[illegible]-25[illegible]-11  Sequence 11, Appl
   44      31    58.5    358  4  US-09-25[illegible]-2[illegible]-31  Sequence 31, Appl
   45      31    58.5    359  6  5474933-8               Patent No. 5474933

                              ALIGNMENTS

RESULT 1
US-08-538-960-7
; Sequence 7, Application US/08538960
; Patent No. 5872230
;  GENERAL INFORMATION:
;    APPLICANT:  Stocco, Douglas M.
;    APPLICANT:  Clark, Barbara J.
;    TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR
;    TITLE OF INVENTION:  REGULATION OF STEROIDOGENESIS
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Akin, Gump, Strauss, Hauer & Feld, L.L.P.
;      STREET:  1900 Frost Bank Plaza, 816 Congress Avenue
;      CITY:  Austin
;      STATE:  TX
;      COUNTRY:  U.S.A.
;      ZIP:  78701
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/538,960
;      FILING DATE:  Concurrently Herewith
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Mayfield, Denise L.
;      REGISTRATION NUMBER:  33,732
;      REFERENCE/DOCKET NUMBER:  [illegible]
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  512/499-6200
;      TELEFAX:  512/499-6290
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-538-960-7

  Query Match             100.0%;  Score 53;  DB 2;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 1.5e+01;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AEHGPTCMV 9
              |||||||||
Db          1 AEHGPTCMV 9
```

```
RESULT 2
US-08-659-254-7
; Sequence 7, Application US/08659[illegible]
; Patent No. 61[illegible]
; GENERAL INFORMATION:
;    APPLICANT: St[illegible] [illegible] M
;    APPLICANT: Clark, [illegible] Barbara [illegible]
;    TITLE OF INVENTION: COMPOSITIONS AND METHODS [illegible]
;    TITLE OF INVENTION: [illegible] GENESIS
;    NUMBER OF SEQUENCES: [illegible]
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE: Akin, Gump, Strauss, Hauer & Feld [illegible]
;      STREET: 1900 Frost Bank Plaza, 816 Congress Avenue
;      CITY: Austin
;      STATE: TX
;      COUNTRY: U.S.A.
;      ZIP: 787[illegible]
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE: Floppy disk
;      COMPUTER: IBM PC compatible
;      OPERATING SYSTEM: PC-DOS/MS-DOS
;      SOFTWARE: PatentIn Release #1.0, Version #[illegible]
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER: US/08/[illegible]
;      FILING DATE: 07-JUN-1996
;      CLASSIFICATION: [illegible]
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER: [illegible]
;      FILING DATE: [illegible] 1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME: Mayfield, Denise L.
;      REGISTRATION NUMBER: [illegible]
;      REFERENCE/DOCKET NUMBER: [illegible]
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE: 512/499-6200
;      TELEFAX: 512/499-[illegible]
;  INFORMATION FOR SEQ ID NO: [illegible]:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH: 9 amino acids
;      TYPE: amino acid
;      STRANDEDNESS: single
;      TOPOLOGY: linear
;    MOLECULE TYPE: peptide
US-08-659-254-7

  Query Match          [illegible]
  Best Local Similarity [illegible]  Pred. No. [illegible]
  Matches [illegible]  Conservative [illegible]  Mismatches [illegible]

Qy          1 AEHGPTCMV 9
              |||||||||
Db          1 AEHGPTCMV 9


RESULT 3
US-08-538-960-[illegible]
; Sequence [illegible], Application US/08538960
; Patent No. 5872[illegible]
; GENERAL INFORMATION:
;    APPLICANT: St[illegible] [illegible] M
;    APPLICANT: Clark, Barbara [illegible]
;    TITLE OF INVENTION: [illegible] MODULATING AND METHODS [illegible]
;    TITLE OF INVENTION: REGULATION OF [illegible]
;    NUMBER OF SEQUENCES: [illegible]
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE: Akin, Gump, Strauss, Hauer & Feld
;      STREET: 1900 Frost Bank Plaza, 816 Congress Avenue
;      CITY: Austin
;      STATE: TX
;      COUNTRY: U.S.A.
```

```
;      ZIP: 78701
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE: Floppy disk
;      COMPUTER: IBM PC compatible
;      OPERATING SYSTEM: PC-DOS/MS-DOS
;      SOFTWARE: PatentIn Release #1.0, Version #[illegible]
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER: US/08/538,960
;      FILING DATE: [illegible]
;      CLASSIFICATION: [illegible]
;    ATTORNEY/AGENT INFORMATION:
;      NAME: Mayfield, Denise [illegible]
;      REGISTRATION NUMBER: [illegible]
;      REFERENCE/DOCKET NUMBER: [illegible]
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE: 512/499-6200
;      TELEFAX: 512/499-[illegible]
;  INFORMATION FOR SEQ ID NO: [illegible]:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH: 14 amino acids
;      TYPE: amino acid
;      STRANDEDNESS: [illegible]
;      TOPOLOGY: [illegible]
;    MOLECULE TYPE: peptide
US-08-538-960-[illegible]

  Query Match          [illegible]  Score [illegible]  Length 14;
  Best Local Similarity [illegible]  Pred. No. [illegible]
  Matches [illegible]  Conservative [illegible]  Mismatches [illegible]  Indels [illegible]  Gaps [illegible]

Qy          1 AE[illegible]CMV [illegible]
              [illegible]
Db          1 AEHGPTCMV [illegible]


RESULT 4
US-08-659-254-[illegible]
; Sequence [illegible], Application US/08659254
; Patent No. [illegible]
; GENERAL INFORMATION:
;    APPLICANT: St[illegible] [illegible] M
;    APPLICANT: Clark, [illegible] Barbara [illegible]
;    TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR REGULATION
;    TITLE OF INVENTION: OF [illegible] GENESIS
;    NUMBER OF SEQUENCES: [illegible]
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE: Akin, Gump, Strauss, Hauer & Feld, L.L.P.
;      STREET: 1900 Frost Bank Plaza, 816 Congress Avenue
;      CITY: Austin
;      STATE: TX
;      COUNTRY: U.S.A.
;      ZIP: 78701
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE: Floppy disk
;      COMPUTER: IBM PC compatible
;      OPERATING SYSTEM: PC-DOS/MS-DOS
;      SOFTWARE: PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER: US/08/659,254
;      FILING DATE: 07-JUN-1996
;      CLASSIFICATION: [illegible]
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER: US/08/[illegible]
;      FILING DATE: [illegible]
;    ATTORNEY/AGENT INFORMATION:
;      NAME: Mayfield, Denise L.
;      REGISTRATION NUMBER: [illegible]
;      REFERENCE/DOCKET NUMBER: [illegible]
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE: 512/499-6200
;      TELEFAX: 512/499-[illegible]
;  INFORMATION FOR SEQ ID NO: [illegible]:
```

```
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  14 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
US-08-659-254-3


  Query Match             100.0%;  Score 5 ;  DB 4;  Length 14;
  Best Local Similarity   100.0%;  Pred. No. 0.0014;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AEHGPTCMV 9
           |||||||||
Db       1 AEHGPTCMV 9


RESULT   5
US-08-659-254-2
; Sequence 2, Application US/08659254
[illegible]tent No. 6194555
[illegible]ENERAL INFORMATION:
    APPLICANT:  Stocco, Douglas M.
;     APPLICANT:  Clark, Dr. Barbara J.
;     TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR REGULATION
;     TITLE OF INVENTION:  OF STEROIDOGENESIS
;     NUMBER OF SEQUENCES:  19
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Akin, Gump, Strauss, Hauer & Feld, L.L.P.
;       STREET:  1900 Frost Bank Plaza, 816 Congress Avenue
;       CITY:  Austin
;       STATE:  TX
;       COUNTRY:  U.S.A.
;       ZIP:  78701
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/659,254
;       FILING DATE:  07-JUN-1996
;       CLASSIFICATION:  530
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/538,960
;       FILING DATE:  04-OCT-1995
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Mayfield, Denise L.
        REGISTRATION NUMBER:  33,732
[illegible]     REFERENCE/DOCKET NUMBER:  43375.0006
      TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  512/499-6200
;       TELEFAX:  512/499-6290
;   INFORMATION FOR SEQ ID NO:  2:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  276 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-659-254-2


  Query Match             100.0%;  Score 5 ;  DB 4;  Length 276;
  Best Local Similarity   100.0%;  Pred. No. 0.027;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AEHGPTCMV 9
           |||||||||
Db     217 AEHGPTCMV 225
```

```
RESULT   6
US-08-538-960-2
; Sequence 2, Application US/08538960
; Patent No. 5872230
;  GENERAL INFORMATION:
;     APPLICANT:  Stocco, Douglas M.
;     APPLICANT:  Clark, Barbara J.
;     TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR
;     TITLE OF INVENTION:  REGULATION OF STEROIDOGENESIS
;     NUMBER OF SEQUENCES:  14
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Akin, Gump, Strauss, Hauer & Feld, L.L.P.
;       STREET:  1900 Frost Bank Plaza, 816 Congress Avenue
;       CITY:  Austin
;       STATE:  TX
;       COUNTRY:  U.S.A.
;       ZIP:  78701
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
,       COMPUTER:  IBM PC compatible
,       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
,       APPLICATION NUMBER:  US/08/538,960
;       FILING DATE:  Concurrently Herewith
;       CLASSIFICATION:  530
,     ATTORNEY/AGENT INFORMATION:
;       NAME:  Mayfield, Denise L.
,       REGISTRATION NUMBER:  33,732
;       REFERENCE/DOCKET NUMBER:  43375-[illegible]
,     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  512/499-6200
;       TELEFAX:  512/499-6290
;   INFORMATION FOR SEQ ID NO:  2:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  284 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-538-960-2


  Query Match             100.0%;  Score 5 ;  DB 2;  Length 284;
  Best Local Similarity   100.0%;  Pred. No. 0.028;
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AEHGPTCMV 9
           |||||||||
Db     217 AEHGPTCMV 225


RESULT   7
US-08-410-540-2
; Sequence 2, Application US/08410540
; Patent No. 5807678
;  GENERAL INFORMATION:
;     APPLICANT:  Miller, Walter L.
;     APPLICANT:  Lin, Dong
;     APPLICANT:  Strauss III, Jerome
;     TITLE OF INVENTION:  IDENTIFICATION OF GENE MUTATIONS
,     TITLE OF INVENTION:  ASSOCIATED WITH CONGENITAL LIPOID ADRENAL HYPERPLASIA
;     NUMBER OF SEQUENCES:  30
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Cooley Godward Castro Huddleson & Tatum
;       STREET:  5 Palo Alto Square
;       CITY:  Palo Alto
;       STATE:  CA
;       COUNTRY:  US
;       ZIP:  94306-2155
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;        OPERATING SYSTEM: PC-DOS/MS-DOS
;        SOFTWARE: PatentIn Release #1.0, Version #1.[illegible]
;      CURRENT APPLICATION DATA:
;        APPLICATION NUMBER: US/08/[illegible]
;        FILING DATE: 23-MAR-[illegible]
;        CLASSIFICATION: [illegible]
;      ATTORNEY/AGENT INFORMATION:
;        NAME: Neeley, Richard L.
;        REGISTRATION NUMBER: 30,[illegible]
;        REFERENCE/DOCKET NUMBER: UCAL [illegible]
;      TELECOMMUNICATION INFORMATION:
;        TELEPHONE: 4[illegible] 8[illegible]
;        TELEFAX: 415-857-[illegible]
;        TELEX: [illegible] CYPA
;    INFORMATION FOR SEQ ID NO: 2:
;      SEQUENCE CHARACTERISTICS:
;        LENGTH: 2[illegible] amino acids
;        TYPE: amino acid
;        TOPOLOGY: linear
;      MOLECULE TYPE: protein
US-08-410-546-2

  Query Match             100.0%;  Score [illegible];  DB [illegible];  Length [illegible];
  Best Local Similarity   100.0%;  Pred. No. [illegible];
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AEHGPICMV 9
                |||||||||
Db        218 AEHGPICMV 226


RESULT 8
US-08-659-254-18
; Sequence 18, Application US/08659254
; Patent No. 61945[illegible]
;  GENERAL INFORMATION:
;    APPLICANT: Stevens, William M
;    APPLICANT: Clark, Dr. Barbara
;    TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODULATION
;    TITLE OF INVENTION: OF STEROIDOGENESIS
;    NUMBER OF SEQUENCES: 19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE: Akin, Gump, Strauss, [illegible]
;      STREET: 1900 Frost Bank Plaza, 816 [illegible]
;      CITY: Austin
;      STATE: TX
;      COUNTRY: U.S.A.
;      ZIP: 78701
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE: Floppy disk
;      COMPUTER: IBM PC compatible
;      OPERATING SYSTEM: PC-DOS/MS-DOS
;      SOFTWARE: PatentIn Release #1.0, Version #1.[illegible]
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER: US/08/[illegible]
;      FILING DATE: [illegible] JUN 1996
;      CLASSIFICATION: [illegible]
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER: US 08/[illegible]
;      FILING DATE: [illegible] OCT 1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME: Mayfield, Denise L.
;      REGISTRATION NUMBER: [illegible]
;      REFERENCE/DOCKET NUMBER: [illegible]
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE: 512/499-[illegible]
;      TELEFAX: 512/4[illegible]
;  INFORMATION FOR SEQ ID NO: 18:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH: 285 amino acids
;      TYPE: amino acid
;      TOPOLOGY: linear
```

```
US-08-659-254-18

  Query Match             [illegible];  Score [illegible];  DB [illegible];  Length 285;
  Best Local Similarity   [illegible];  Pred. No. [illegible];
  Matches [illegible];  Conservative [illegible];  Mismatches [illegible];  Indels [illegible];  Gaps [illegible];

Qy          1 AEHGPICMV 9

Db        [illegible] AEHGPICMV [illegible]


RESULT 9
US-08-[illegible]
; Sequence [illegible], Application US/08[illegible]
; Patent No. [illegible]
;  GENERAL INFORMATION:
;    APPLICANT: [illegible]
;    APPLICANT: [illegible]
;    APPLICANT: [illegible], Robert [illegible]
;    APPLICANT: Xu, [illegible]
;    APPLICANT: Ma[illegible]
;    TITLE OF INVENTION: HELIX PROTEINS, NUCLEIC ACIDS, AND
;    TITLE OF INVENTION: ANTIBODIES, AND RELATED METHODS AND COMPOSITIONS
;    NUMBER OF SEQUENCES: [illegible]
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE: PENNIE & EDMONDS
;      STREET: [illegible] Avenue of the Americas
;      CITY: New York
;      STATE: New York
;      COUNTRY: U.S.A.
;      ZIP: [illegible]
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE: Floppy disk
;      COMPUTER: IBM PC compatible
;      OPERATING SYSTEM: PC-DOS/MS-DOS
;      SOFTWARE: PatentIn Release #[illegible] Version #[illegible]
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER: US/08/[illegible]
;      FILING DATE: [illegible] JAN [illegible]
;      CLASSIFICATION: [illegible]
;    ATTORNEY/AGENT INFORMATION:
;      NAME: Misrock, S. Leslie
;      REGISTRATION NUMBER: [illegible]
;      REFERENCE/DOCKET NUMBER: [illegible]
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE: (212) [illegible]
;      TELEFAX: (212) [illegible]
;      TELEX: [illegible] PENNIE
;  INFORMATION FOR SEQ ID NO: [illegible]
;    SEQUENCE CHARACTERISTICS:
;      LENGTH: [illegible] amino acids
;      TYPE: amino acid
;      TOPOLOGY: unknown
;    MOLECULE TYPE: protein
US-08-[illegible]

  Query Match             [illegible];  Score [illegible];  DB [illegible];  Length 2[illegible];
  Best Local Similarity   [illegible];  Pred. No. [illegible];
  Matches [illegible];  Conservative [illegible];  Mismatches [illegible];  Indels [illegible];  Gaps [illegible];

Qy          [illegible] HG[illegible]

Db        [illegible] HG[illegible]


RESULT 1[illegible]
US-08-[illegible]A
; Sequence [illegible], Application US/08[illegible]A
; Patent No. [illegible]
;  GENERAL INFORMATION:
;    APPLICANT: [illegible]
;    TITLE OF INVENTION:  Therapeutic And Diagnostic Methods
;    TITLE OF INVENTION:  And Compositions Based On No. 5786158; Proteins And
;    TITLE OF INVENTION:  Nucleic Acids
;    NUMBER OF SEQUENCES:  21
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/083,590A
;      FILING DATE:  25-JUN-1993
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
       NAME:  Misrock, S. Leslie
       REGISTRATION NUMBER:  18,872
       REFERENCE/DOCKET NUMBER:  7326-015
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212 790-9090
;      TELEFAX:  212 8698864/9741
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  20:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2556 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  peptide
US-08-083-590A-20


  Query Match             66.0%;  Score 3 ;  DB 1;  Length 2556;
  Best Local Similarity   100.0%;  Pred. No. 3.7e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3 HGPTC 7
           |||||
Db     171 HGPTC 175


RESULT  11
US-08-532-384-20
; Sequence 20, Application US/08532384
  tent No. 6083904
  ENERAL INFORMATION:
;    APPLICANT:  Artavanis-Tsakonas, S. et al.
;    TITLE OF INVENTION:  Therapeutic And Diagnostic Methods
;    TITLE OF INVENTION:  And Compositions Based On No. 6083904; Proteins And
;    TITLE OF INVENTION:  Nucleic Acids
;    NUMBER OF SEQUENCES:  21
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/532,384
;      FILING DATE:
;      CLASSIFICATION:  424
```

```
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/083,590
;      FILING DATE:  25-JUN-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Misrock, S. Leslie
;      REGISTRATION NUMBER:  18,872
;      REFERENCE/DOCKET NUMBER:  7326-015
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212 790-9090
;      TELEFAX:  212 8698864/9741
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  20:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2556 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  unknown
,    MOLECULE TYPE:  peptide
US-08-532-384-20

  Query Match             66.0%;  Score 3 ;  DB [illegible];  Length 2556;
  Best Local Similarity   100.0%;  Pred. No. 3.7e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3 HGPTC 7
           |||||
Db     171 HGPTC 175


RESULT  12
US-08-173-510B-40
; Sequence 40, Application US/08173510B
; Patent No. 5747296
;  GENERAL INFORMATION:
;    APPLICANT:  MATTHEW MOYLE, ET AL.
;    TITLE OF INVENTION:  NOVEL NEUTROPHIL INHIBITORS
;    NUMBER OF SEQUENCES:  104
,    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Lyon & Lyon
;      STREET:  633 West Fifth Street
;      STREET:  Suite 4700
;      CITY:  Los Angeles
;      STATE:  California
;      COUNTRY:  U.S.A.
;      ZIP:  90071
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;      MEDIUM TYPE:  storage
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  IBM P.C. DOS 5.0
;      SOFTWARE:  Word Perfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/173,510B
;      FILING DATE:  23-DEC-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/151,063
;      FILING DATE:  10-NOV-1993
;      APPLICATION NUMBER:  08/060,433
;      FILING DATE:  11-MAY-1993
;      APPLICATION NUMBER:  07/996,972
;      FILING DATE:  24-DEC-1992
;      APPLICATION NUMBER:  07/881,721
;      FILING DATE:  11-MAY-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BIGGS, SUZANNE L.
;      REGISTRATION NUMBER:  30,158
;      REFERENCE/DOCKET NUMBER:  [illegible]
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (213) 489-1600
;      TELEFAX:  (213) 955-0440
;      TELEX:  67-3510
;  INFORMATION FOR SEQ ID NO:  40:
```

```
;    SEQUENCE CHARACTERISTICS:
;      LENGTH: [illegible] AMINO ACIDS
;      TYPE: AMINO ACID
;      TOPOLOGY: LINEAR
;    MOLECULE TYPE: PEPTIDE
US-08-173-510B-4

  Query Match            [illegible]
  Best Local Similarity  [illegible]
  Matches    5;  Conservative  [illegible]  Mismatches  [illegible]

Qy          2 EHGPT[illegible]
              || |||
Db          2 EHDPTC[illegible]


[illegible]LT 13
[illegible]8-458-218-40
[illegible]equence 40, Application US/08458218
; Patent No. 578[illegible]
;  GENERAL INFORMATION:
;    APPLICANT: MATTHEW MOYLE ET A[illegible]
;    TITLE OF INVENTION: NOVEL NEUTROPHIL INHIBITORS
;    NUMBER OF SEQUENCES: [illegible]
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE: Lyon & Lyon
;      STREET: 633 West Fifth Street
;      STREET: Suite 47[illegible]
;      CITY: Los Angeles
;      STATE: California
;      COUNTRY: U.S.A.
;      ZIP: 90071
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
;      MEDIUM TYPE: Storage
;      COMPUTER: IBM Compatible
;      OPERATING SYSTEM: [illegible]
;      SOFTWARE: Word Perfect [illegible]
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER: US/08/4[illegible]
;      FILING DATE:
;      CLASSIFICATION: 4[illegible]
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER: [illegible]
;      FILING DATE: [illegible] NOVEMBER 199[illegible]
;      APPLICATION NUMBER: 08/060,433
;      FILING DATE: [illegible] MAY 1993
;      APPLICATION NUMBER: 07/881,[illegible]
;      FILING DATE: [illegible] MAY 1992
;      APPLICATION NUMBER: [illegible]
;      FILING DATE: [illegible] DEC [illegible]
;    ATTORNEY/AGENT INFORMATION:
;      NAME: BIGGS, SUZANNE L.
;      REGISTRATION NUMBER: [illegible]
;      REFERENCE/DOCKET NUMBER: [illegible]
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE: ([illegible]) 4[illegible]
;      TELEFAX: (213) 955-[illegible]
;      TELEX: 67[illegible]
;  INFORMATION FOR SEQ ID NO: [illegible]:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH: 9 AMINO ACIDS
;      TYPE: AMINO ACID
;      TOPOLOGY: LINEAR
;    MOLECULE TYPE: PEPTIDE
;    FEATURE:
;      OTHER INFORMATION:
US-08-458-218-40

  Query Match            [illegible]
  Best Local Similarity  [illegible]  Pred. No. [illegible]
```

```
  Matches  [illegible]  Conservative  [illegible]  Mismatches  [illegible]  Indels  [illegible]  Gaps  [illegible]

Qy          [illegible]

Db          [illegible]


RESULT 14
US-08-45[illegible]
; Sequence [illegible], Application US/08[illegible]
; Patent No. [illegible]
;  GENERAL INFORMATION:
;    APPLICANT: MATTHEW MOYLE ET AL.
;    TITLE OF INVENTION: NOVEL NEUTROPHIL INHIBITORS
;    NUMBER OF SEQUENCES: [illegible]
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE: [illegible]
;      STREET: [illegible] West Fifth Street
;      STREET: Suite 47[illegible]
;      CITY: [illegible]
;      STATE: [illegible]
;      COUNTRY: U.S.A.
;      ZIP: [illegible]
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE: [illegible] Diskette, 1.44 Mb
;      MEDIUM TYPE: [illegible]
;      COMPUTER: IBM [illegible]
;      OPERATING SYSTEM: [illegible]
;      SOFTWARE: Word Perfect [illegible]
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER: [illegible]
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER: [illegible]
;      FILING DATE: [illegible]
;      APPLICATION NUMBER: [illegible]
;      FILING DATE: [illegible] NOV [illegible]
;      APPLICATION NUMBER: [illegible]
;      FILING DATE: [illegible] MAY [illegible]
;      APPLICATION NUMBER: [illegible]
;      FILING DATE: [illegible] DEC [illegible]
;      APPLICATION NUMBER: [illegible]
;      FILING DATE: [illegible] MAY [illegible]
;    ATTORNEY/AGENT INFORMATION:
;      NAME: [illegible], SUZANNE [illegible]
;      REGISTRATION NUMBER: [illegible]
;      REFERENCE/DOCKET NUMBER: [illegible]
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE: ([illegible]) 4[illegible]
;      TELEFAX: ([illegible]) [illegible]
;      TELEX: [illegible]
;  INFORMATION FOR SEQ ID NO: [illegible]:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH: [illegible] AMINO ACIDS
;      TYPE: AMINO ACID
;      TOPOLOGY: LINEAR
;    MOLECULE TYPE: PEPTIDE
US-0[illegible]

  Query Match            [illegible]  Length 9[illegible]
  Best Local Similarity  [illegible]  Pred. No. [illegible]
  Matches  [illegible]  Conservative  [illegible]  Mismatches  [illegible]  Indels  [illegible]  Gaps  [illegible]

Qy          [illegible]

Db          [illegible]


RESULT 15
US-0[illegible]
; Sequence [illegible], Application US/[illegible]
; Patent No. [illegible]
;  GENERAL INFORMATION:
;    APPLICANT:  WANG, Xun
;    APPLICANT:  DUVICK, Jonathan P.
;    APPLICANT:  BRIGGS, Steven P.
;    TITLE OF INVENTION:  PCR BASED CDNA SUBTRACTIVE CLONING
;    TITLE OF INVENTION:  METHOD
;    NUMBER OF SEQUENCES:  39
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  3000 K Street, N.W., Suite [illegible]
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20007-5109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/858,767
;      FILING DATE:  19-MAY-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/481,687
;      FILING DATE:  07-JUN-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BENT, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  33229/325[illegible]
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)672-5300
;      TELEFAX:  (202)672-5399
;      TELEX:  904136
;  INFORMATION FOR SEQ ID NO:  28:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  99 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-858-767-28


  Query Match             62.3%;  Score 3 ;  DB 2;  Length 99;
  Best Local Similarity   44.4%;  Pred. No. 3;
  Matches    4;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 AEHGPTCMV 9
           :|:|| |::
        28 SENGPVCVI 36


Search completed: August 22, 2001, 14 39:12
Job time: 366 sec
```

```
                         GenCore version 4.5
                  Copyright (c) 1993 - 200[illegible] Compugen Ltd.

OM protein - protein search, using sw model

Run on:          August 22, 2001, 14:29:10 ; Search time 101.96 Seconds
                                             (without alignments)
                                             3 568 Million cell updates/sec

Title:           US-08-659-254-6
Perfect score:   38
Sequence:        1 NQEGWK 6

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0 5

Searched:        412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters:        412676

[illegible]imum DB seq length: 0
[illegible]imum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       A_Geneseq_0601:*
               1:  /SIDS8/gcgdata/geneseq/geneseqp/AA19[illegible].DAT:*
               2:  /SIDS8/gcgdata/geneseq/geneseqp/AA1[illegible].DAT:*
               3:  /SIDS8/gcgdata/geneseq/geneseqp/AA198[illegible].DAT:*
               4:  /SIDS8/gcgdata/geneseq/geneseqp/AA198[illegible].DAT:*
               5:  /SIDS8/gcgdata/geneseq/geneseqp/AA1984.DAT:*
               6:  /SIDS8/gcgdata/geneseq/geneseqp/AA1985.DAT:*
               7:  /SIDS8/gcgdata/geneseq/geneseqp/AA1986.DAT:*
               8:  /SIDS8/gcgdata/geneseq/geneseqp/AA1987.DAT:*
               9:  /SIDS8/gcgdata/geneseq/geneseqp/AA1988.DAT:*
               10:  /SIDS8/gcgdata/geneseq/geneseqp/AA1989.DAT:*
               11:  /SIDS8/gcgdata/geneseq/geneseqp/AA1990.DAT:*
               12:  /SIDS8/gcgdata/geneseq/geneseqp/AA199[illegible].DAT:*
               13:  /SIDS8/gcgdata/geneseq/geneseqp/AA1992.DAT:*
               14:  /SIDS8/gcgdata/geneseq/geneseqp/AA1993.DAT:*
               15:  /SIDS8/gcgdata/geneseq/geneseqp/AA1994.DAT:*
               16:  /SIDS8/gcgdata/geneseq/geneseqp/AA199[illegible].DAT:*
               17:  /SIDS8/gcgdata/geneseq/geneseqp/AA1996.DAT:*
               18:  /SIDS8/gcgdata/geneseq/geneseqp/AA1997.DAT:*
               19:  /SIDS8/gcgdata/geneseq/geneseqp/AA1998.DAT:*
               20:  /SIDS8/gcgdata/geneseq/geneseqp/AA199[illegible].DAT:*
               21:  /SIDS8/gcgdata/geneseq/geneseqp/AA20[illegible].DAT:*
               22  /SIDS8/gcgdata/geneseq/geneseqp/AA20[illegible].DAT:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed
      and is derived by analysis of the total score distribution.

                                    SUMMARIES

                    %
Result              Query
   No.    Score   Match Length DB   ID                          Description
------------------------------------------------------------------------------
     1       38   100.0      6  22  AAY728[illegible]          Steroidogenic acut
     2       38   100.0     11  22  AAY728[illegible]          Mouse steroidogeni
     3       38   100.0     12  20  AAW74[illegible]           Mouse steroidogeni
     4       38   100.0     12  22  AAY728[illegible]          Mouse tryptic pept
     5       38   100.0    284  17  AAW132[illegible]          Human steroidogene
     6       38   100.0    284  20  AAW74[illegible]           Mouse steroidogeni
     7       38   100.0    284  22  AAY728[illegible]          Mouse steroidogeni
     8       38   100.0    285  17  AAW132[illegible]          Human steroidogene
     9       38   100.0    285  17  AAW13[illegible]           Human steroid acute
    10       38   100.0    285  17  AAW00814                   Human steroidogene
    11       38   100.0    285  17  AAW132[illegible]          Human steroidogene
    12       38   100.0    285  17  AAW13[illegible]           Human steroidogene
    13       38   100.0    285  17  AAW13[illegible]           Human steroidogene
    14       38   100.0    285  22  AAY72[illegible]           Human steroidogeni
    15       35    92.1    102  20  AAY28[illegible]           pk8[illegible]5_1 secreted p
    16       35    92.1    824  21  AAB23[illegible]           Bacillus sp. KSM-S
    17       33    86.8     61  21  AAG[illegible]             Arabidopsis thalia
    18       33    86.8    514  12  AAR15[illegible]           Fusarium oxysporum
    19       33    86.8   3587  14  AAR34[illegible]           Bacillus subtilis
    20       33    86.8   3588  14  AAR34[illegible]           Bacillus subtilis
    21       32    84.2     53  21  AAY64[illegible]           Human 5' EST relat
    22       32    84.2     85  20  AAY7421[illegible]         Human prostate tum
    23       32    84.2     96  15  AAR46[illegible]           Human cathepsin L
    24       32    84.2     96  20  AAY423[illegible]          Human cathepsin L
    25       32    84.2    206  21  AAY[illegible]             Human cancer assoc
    26       32    84.2    219  20  AAW9[illegible]            Expressed antigen
    27       32    84.2    271  19  AAW9818[illegible]         H. pylori GHP0 958
    28       32    84.2    286  21  AAY321[illegible]          Plant retroelement
    29       32    84.2    333   9  AAP82[illegible]          pHu 16 sequence en
    30       32    84.2    333  13  AAR28[illegible]           Human procathepsin
    31       32    84.2    333  19  AAW47[illegible]           Human procathepsin
    32       32    84.2    334  20  AAY02[illegible]           Polypeptide identi
    33       32    84.2    334  20  AAW94[illegible]           Human cathepsin (L
    34       32    84.2    334  21  AAY8[illegible]            Human cathepsin L2
    35       32    84.2    616  10  AAP9[illegible]            Sequence of Newcas
    36       32    84.2    616  11  AAR06[illegible]           Newcastle disease
    37       32    84.2    877  21  AAB[illegible]             E. coli proliferat
    38       32    84.2   1801  21  AAY3[illegible]            Plant generic retr
    39       32    84.2   1974  19  AAW98[illegible]           H. pylori GHP0 57
    40       32    84.2   2115  19  AAW5[illegible]            Rubella virus RA27
    41       32    84.2   2205  16  AAR7[illegible]            Infect ous rubella
    42       31    81.6    [illegible]  21  AAG2[illegible]            Arabidopsis thalia
    43       31    81.6    [illegible]  21  AA[illegible]              A. thaliana enviro
    44       31    81.6    112  21  AAG2[illegible]            Arabidopsis thalia
    45       30    [illegible]      7  [illegible]  AAW[illegible]             [illegible]

                                    ALIGNMENTS

RESULT    1
AAY72835
ID   AAY72835 standard; peptide; 6 AA.
XX
AC   AAY72835;
XX
DT   [illegible]-MAY 2001  (first entry)
XX
DE   Steroidogenic acute regulatory (StAR) peptide #
XX
KW   Steroidogenic acute regulatory protein; StAR;
KW   gene therapy; steroid hormone-dependent disorder; LCAH;
KW   lipoid congenital adrenal hyperplasia; sexual maturation;
KW   infertility; androgen-responsive tumour; precocious puberty;
KW   hypogonadotropic hypogonadism; McCune-Albright syndrome;
KW   adrenal-hypoplasia congenita; cytostatic.
XX
OS   Unidentified.
XX
PN   US6194555-B1
XX
PD   27-FEB-2001.
XX
PF   07-JUN-1996;    96US-0659254.
XX
PR   04-NOV-1995;    95US-0538960.
XX
PA   (UYTE-) UNIV TEXAS TECH HEALTH SCI CENT.
XX
PI   Stocco DM,  Clark BJ;
XX
DR   WPI; 2001-[illegible]/22.
XX
PT   Novel nucleic acid encoding steroidogenic acute regulatory protein,
PI   useful for treating steroid hormone dependent disorders
XX
PS   Disclosure; Column [illegible]; 41pp; English.
XX
CC   The invention relates to steroidogenic acute regulatory protein
CC   (STAR) and nucleic acid molecules encoding them. STAR proteins
CC   are used for regulating steroidogenesis. More particularly, STAR
CC   proteins are involved in regulating cholesterol transport into
CC   the mitochondria of a cell. The nucleic acid molecules of the
CC   invention are useful in gene therapy for the treatment of steroid
CC   hormone dependent disorders, particularly lipoid congenital
CC   adrenal hyperplasia (LCAH), infertility, sexual maturation,
CC   androgen-responsive tumours, precocious puberty, McCune-Albright
CC   syndrome, adrenal hypoplasia congenita and hypogonadotropic
CC   hypogonadism. They are also useful for recombinant preparation
CC   of the steroidogenic acute regulatory protein.
CC   The present sequence is steroidogenic acute regulatory (STAR)
CC   peptide. This peptide is useful for designing [illegible]
CC   screening and identifying antigenic determinants of STAR protein.
XX
SQ   Sequence   [illegible] AA;

  Query Match             100.0%;  Score [illegible];  DB [illegible];  Length [illegible];
  Best Local Similarity   [illegible]%;  Pred. No. [illegible];
  Matches    6;  Conservative    [illegible];  Mismatches    [illegible];  Indels    [illegible];  Gaps    [illegible];

Qy          1 NLEGWK 6
              ||||||
Db          1 nlegwk 6


RESULT 2
AAY72837
ID   AAY72837 standard; peptide; [illegible] AA.
XX
AC   AAY72837;
XX
DT   31-MAY-2001  (first entry)
XX
DE   Mouse steroidogenic acute regulatory (STAR) peptide #[illegible].
XX
KW   Mouse; steroidogenic acute regulatory protein; STAR;
KW   gene therapy; steroid hormone dependent disorder; [illegible]
KW   lipoid congenital adrenal hyperplasia; sexual maturation;
KW   infertility; androgen responsive tumour; precocious puberty;
KW   hypogonadotropic hypogonadism; McCune Albright syndrome;
KW   adrenal hypoplasia congenita; [illegible]
XX
OS   Mus sp.
XX
PN   US6194555-B1.
XX
PD   27-FEB-2001.
XX
PF   07-JUN-1996;  [illegible]
XX
PR   04-NOV-1995;  [illegible]
XX
PA   (UYTE ) UNIV TEXAS [illegible] HEALTH [illegible]
XX
PI   Stocco DM,  Clark BJ;
XX
DR   WPI; 2001-217[illegible]
XX
PT   Novel nucleic acid encoding steroidogenic acute regulatory protein,
PT   useful for treating steroid hormone dependent disorders
XX
PS   Disclosure; Column 51; 41pp; English.
XX
CC   The invention relates to steroidogenic acute regulatory protein
CC   (STAR) and nucleic acid molecules encoding them. STAR proteins
CC   are used for regulating steroidogenesis. More particularly, STAR
CC   proteins are involved in regulating cholesterol transport into
CC   the mitochondria of a cell. The nucleic acid molecules of the
CC   invention are useful in gene therapy for the treatment of steroid
CC   hormone dependent disorders, particularly lipoid congenital
CC   adrenal hyperplasia (LCAH), infertility, sexual maturation,
CC   androgen-responsive tumours, precocious puberty, McCune-Albright
CC   syndrome, adrenal hypoplasia congenita and hypogonadotropic
CC   hypogonadism. They are also useful for recombinant preparation
CC   of the steroidogenic acute regulatory protein.
CC   The present sequence is steroidogenic acute regulatory (STAR)
CC   peptide. This peptide [illegible]
CC   STAR protein. This peptide is useful for designing [illegible]
CC   for screening and identifying antigenic determinants of STAR protein.
XX
SQ   Sequence   11 AA;

  Query Match             [illegible];  Score [illegible];  DB [illegible];  Length [illegible];
  Best Local Similarity   [illegible];  Pred. No. [illegible];
  Matches    [illegible];  Conservative    [illegible];  Mismatches    [illegible];  Indels    [illegible];  Gaps    [illegible];

Qy          1 NLEGWK 6

Db          1 [illegible]wk [illegible]


RESULT [illegible]
AAW[illegible]
ID   AAW[illegible] standard; [illegible] AA.
XX
AC   AAW[illegible];
XX
DT   [illegible]-MAY-[illegible]  (first entry)
XX
DE   Mouse steroidogenic acute regulatory protein tryptic peptide.
XX
KW   STAR; steroidogenic acute regulatory protein; McCune-Albright syndrome;
KW   steroid hormone dependent disorder; adrenal hyperplasia congenita; LCAH;
KW   lipoid congenital adrenal hyperplasia; infertility; sexual maturation;
KW   androgen responsive tumour; hypogonadism; [illegible] precocious puberty.
XX
OS   Mus sp.
XX
PN   [illegible]-A.
XX
PD   [illegible]
XX
PF   [illegible]-NOV-[illegible]
XX
PR   [illegible]-NOV-[illegible]
XX
PA   (UYTE ) UNIV TEXAS [illegible] HEALTH SCI CENT.
XX
PI   Clark BJ,  Stocco DM;
XX
DR   WPI; [illegible]
XX
PT   New nucleic acid encoding steroidogenic acute regulatory protein
PT   useful for treatment of steroid hormone dependent disorders such as
PT   lipoid congenital adrenal hyperplasia, infertility and
PT   androgen-responsive tumours
XX
PS   Example [illegible]; [illegible]pp; English.
XX
CC   This sequence is a tryptic peptide fragment of the steroidogenic acute
CC   regulatory protein (STAR) of the invention. The STAR coding sequence and
CC   [illegible] can be used as probes and primers for [illegible]
CC   hybridisation studies and in the treatment of steroid hormone dependent
CC   disorders such as lipoid congenital adrenal hyperplasia (LCAH),
CC   infertility, sexual maturation, androgen-responsive tumours, precocious
CC   puberty, McCune-Albright syndrome, adrenal hypoplasia congenita and
CC   hypogonadism. Measurement levels of STAR may also indicate reduced
CC   [illegible]
```

```
XX
SQ   Sequence   12 AA;


  Query Match             100.0%;  Score 38;  DB 20;  Length 12;
  Best Local Similarity   100.0%;  Pred. No. .41;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 NQEGWK 6
              ||||||
Db          7 nqeqwk 12


RESULT    4
AAY72833
ID   AAY72833 standard; peptide; 12 AA.
XX
AC   AAY72833;
XX
DT   31-MAY-2001  (first entry)
XX
DE   Mouse tryptic peptide #25 for analysis of 30kDa StAR protein.
XX
KW   Mouse; steroidogenic acute regulatory protein; StAR;
KW   gene therapy; steroid hormone-dependent disorder; LCAH;
KW   lipoid congenital adrenal hyperplasia; sexual maturation;
KW   infertility; androgen-responsive tumour; precocious puberty;
KW   hypogonadotropic hypogonadism; McCune Albright syndrome;
KW   adrenal-hypoplasia congenita; cytostatic; tryptic peptide.
XX
OS   Mus sp.
XX
PN   US6194555-B1.
XX
PD   27 FEB-2001.
XX
PF   07-JUN-1996;   96US-0659254.
XX
PR   04-NOV-1995;   95US-0538960.
XX
PA   (UYTE-) UNIV TEXAS TECH HEALTH SCI CENT.
XX
PI   Stocco DM,  Clark BJ;
XX
DR   WPI; 2001-217936/22.
XX
PT   Novel nucleic acid encoding steroidogenic acute regulatory protein,
PT   useful for treating steroid hormone-dependent disorders -
XX
PS   Example 2; Column 28; 41pp; English.
XX
CC   The invention relates to steroidogenic acute regulatory protein
CC   (StAR) and nucleic acid molecules encoding them. StAR proteins
CC   are used for regulating steroidogenesis. More particularly these
CC   proteins are involved in regulating cholesterol transport into
CC   the mitochondria of a cell. The nucleic acid molecules of the
CC   invention are useful in gene therapy for the treatment of steroid
CC   hormone dependent disorders, particularly, lipoid congenital
CC   adrenal hyperplasia (LCAH), infertility, sexual maturation,
CC   androgen-responsive tumours, precocious puberty, McCune Albright
CC   syndrome, adrenal hypoplasia congenita or hypogonadotropic
CC   hypogonadism. They are also useful for recombinant preparation
CC   of the steroidogenic acute regulatory protein.
CC   The present sequence is tryptic peptide #25. This peptide is
CC   used in the cloning and sequence analysis of cDNA encoding the
CC   30kDa form of StAR protein. This peptide is also used in the
CC   analysis of the StAR protein sequence. Peptide #25 is useful
CC   for designing oligonucleotides for screening and identifying
CC   antigenic determinants of StAR protein.
XX
SQ   Sequence   12 AA;
```

```
  Query Match             100.0%;  Score 38;  DB 22;  Length 12;
  Best Local Similarity   100.0%;  Pred. No. .42;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 NQEGWK 6
              ||||||
Db          7 nqeqwk 12


RESULT    5
AAW13238
ID   AAW13238 standard; Protein; 284 AA
XX
AC   AAW13238;
XX
DT   21 MAY 1997  (first entry)
XX
DE   Human steroidogenesis acute regulatory protein mutant delta Arg272.
XX
KW   Human; steroidogenesis; acute regulatory protein; hStAR; analysis;
KW   mutation; detection; prenatal; genetic defect; congenital; protein;
KW   lipoid adrenal hyperplasia; treatment; prevention; gene;
KW   replacement therapy; hypercholesterolaemia; mutant.
XX
OS   Homo sapiens.
XX
PN   WO9629338 A1.
XX
PD   26-SEP-1996.
XX
PF   22-MAR-1996;   96WO-US03896.
XX
PR   23-MAR-1995;   95US-0410540.
XX
PA   (REGC ) UNIV CALIFORNIA.
PA   (UYPE-) UNIV PENNSYLVANIA.
XX
PI   Lin D,  Miller WL,  Strauss JF;
XX
DR   WPI; 1996-443159/44.
XX
PT   Isolated human steroidogenesis acute regulatory protein gene - used
PT   for detection of mutation(s) of this gene that cause congenital
PT   lipoid adrenal hyperplasia
XX
PS   Claim 16; Page  ; 89pp; English.
XX
CC   The present sequence is the human steroidogenesis acute regulatory
CC   protein (hStAR), congenital lipoid adrenal hyperplasia (CAH)
CC   associated mutant, delta Arg272 (i.e. wild type residue Arg272 has
CC   been deleted). Analysing the hStAR gene for such a mutation is
CC   useful in the detection (e.g. prenatally) of CAH associated
CC   genetic defects, or its transmission to children. CAH can be
CC   treated by protein or gene replacement therapy, which can also be
CC   used to prevent or treat hypercholesterolaemia.
XX
SQ   Sequence   284 AA;


  Query Match             100.0%;  Score 38;  DB 17;  Length 284;
  Best Local Similarity   100.0%;  Pred. No. 5.4;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 NQEGWK 6
              ||||||
Db         92 nqeqwk 97


RESULT    6
AAW74153
ID   AAW74153 standard; Protein; 284 AA.
XX
AC   AAW74153;
```

```
XX
DT   05-MAY-1999  (first entry)
XX
DE   Mouse steroidogenic acute regulatory protein.
XX
KW   StAR; steroidogenic acute regulatory protein; [illegible]
KW   steroid hormone dependent disorder; adrenal hypoplasia [illegible]
KW   lipoid congenital adrenal hyperplasia; infertility; [illegible]
KW   androgen responsive tumour; [illegible]
XX
OS   Mus sp.
XX
PN   US5872231-A.
XX
PD   16-FEB-1999.
XX
PF   04-NOV-1995;   [illegible]
XX
PR   04-NOV-1995;   [illegible]
XX
PA   (UYTE ) UNIV TEXAS [illegible] HEALTH [illegible] CENT
XX
PI   Clark B[illegible] DM
XX
DR   WPI; 1999-169[illegible]/14.
DR   N-PSDB; AAX18[illegible]  AAX[illegible]
XX
PT   New nucleic acid encoding steroidogenic acute regulatory protein
PT   useful for treatment of steroid hormone dependent [illegible]
PT   lipoid congenital adrenal hyperplasia, infertility, [illegible]
PT   androgen responsive tumours.
XX
PS   Claim 1; Col [illegible]; [illegible]
XX
CC   This sequence is the steroidogenic acute regulatory protein (StAR)
CC   of the invention. The StAR coding sequence and [illegible]
CC   derivatives may be used as probes and primers [illegible]
CC   studies and in the treatment of steroid hormone dependent [illegible]
CC   as lipoid congenital adrenal hyperplasia (lCAH), [illegible]
CC   maturation, androgen responsive tumours, [illegible]
CC   McCune Albright syndrome, adrenal hypoplasia [illegible]
CC   Monitoring levels of StAR may [illegible] indicate [illegible]
CC   and so a risk of spontaneous abortion.
XX
SQ   Sequence   28[illegible] AA;

  Query Match              [illegible]
  Best Local Similarity    [illegible]  Pred. No. [illegible]
  Matches [illegible]  Conservative [illegible]  Mismatches [illegible]  Indels [illegible]  Gaps [illegible]

Qy          1 NQESWK [illegible]
              ||||||
Db         93 nqeswk [illegible]


RESULT 7
AAY728[illegible]1
ID   AAY728[illegible]1 standard; Protein; [illegible] AA.
XX
AC   AAY728[illegible];
XX
DT   31-MAY-2000  (first entry)
XX
DE   Mouse steroidogenic acute regulatory (StAR) protein.
XX
KW   Mouse; steroidogenic acute regulatory protein; StAR;
KW   gene therapy; steroid hormone dependent disorder; lCAH;
KW   lipoid congenital adrenal hyperplasia; sexual maturation;
KW   infertility; androgen responsive tumour; [illegible]
KW   hypogonadotropic hypogonadism; McCune Albright syndrome;
KW   adrenal hypoplasia congenita; [illegible]
XX
```

```
OS   Mus sp.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..[illegible]
FT                   /label= Signal peptide
FT   Cleavage-site   [illegible]
FT                   /note= "Site for mitochondrial proteases,
FT                   [illegible]
FT                   intermediate processing peptide"
FT   Protein         [illegible]
FT                   /label= Mature steroidogenic acute regulatory protein
FT   Cleavage-site   [illegible]
FT                   /note= "Site for mitochondrial proteases,
FT                   [illegible]
FT                   intermediate processing peptide"
FT   Modified-site   [illegible]
FT                   /note= "PKA phosphorylation site"
FT   Modified-site   [illegible]
FT                   /note= "PK phosphorylation site"
FT   Region          [illegible]
FT                   /label= Tryptic peptide #1
FT   Region          [illegible]
FT                   /note= "Used to [illegible] antibodies
FT                   against StAR protein"
FT   Region          [illegible]
FT                   /label= Tryptic peptide #2
FT   Modified-site   [illegible]
FT                   /note= "PKA phosphorylation site"
FT   Region          [illegible]
FT                   /label= Tryptic peptide #3
FT   Modified-site   [illegible]
FT                   /note= "Casein kinase (II) phosphorylation site"
FT   Modified-site   [illegible]
FT                   /note= "PKC phosphorylation site"
XX
PN   [illegible]
XX
PD   [illegible]
XX
PF   [illegible]
XX
PR   [illegible] NOV [illegible]
XX
PA   [illegible] UNIV TEXAS [illegible] HEALTH [illegible] CENT
XX
PI   [illegible] DM, Clark [illegible]
XX
DR   WPI; [illegible]
DR   N-PSDB; AA[illegible]
XX
PT   Novel [illegible] steroidogenic acute regulatory protein
PT   useful for treatment of steroid hormone dependent disorders
XX
PS   [illegible]
XX
CC   [illegible] steroidogenic acute regulatory protein
CC   (StAR) and nucleic acid molecules encoding them. StAR proteins
CC   are involved in regulating steroidogenesis. More particularly these
CC   proteins are involved in regulating cholesterol transport [illegible]
CC   [illegible] the nucleic acid molecules of the
CC   invention are useful [illegible] gene therapy for the treatment of steroid
CC   hormone dependent disorders, particularly lipoid congenital
CC   adrenal hyperplasia (lCAH), infertility, sexual maturation,
CC   androgen responsive tumours, [illegible], McCune Albright
CC   syndrome, adrenal hypoplasia congenita or hypogonadotropic
CC   [illegible]. They are also useful for recombinant preparation
CC   of the steroidogenic acute regulatory protein.
CC   The present sequence [illegible] steroidogenic acute regulatory
CC   protein (StAR).
XX
SQ   Sequence   [illegible] AA;
```

```
  Query Match             100.0%;  Score 38;  DB 22;  Length 284;
  Best Local Similarity   100.0%;  Pred. No. 9.4;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 NQEGWK 6
            ||||||
Db       91 nqeqwk 96


RESULT    8
AAW13236
ID   AAW13236 standard; Protein; 285 AA.
XX
AC   AAW13236;
XX
DT   21-MAY-1997  (first entry)
XX
DE   Human steroidogenesis acute regulatory protein mutant Ala218Val.
XX
KW   Human; steroidogenesis; acute regulatory protein; hStAR; analysis;
KW   mutation; detection; prenatal; genetic defect; congenital; protein;
KW   lipoid adrenal hyperplasia; treatment; prevention; gene;
     replacement therapy; hypercholesterolaemia; mutant.

OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc difference 218
FT                   /note= "wild type Ala substituted with Val"
XX
PN   WO9629338-A1.
XX
PD   26-SEP-1996.
XX
PF   22-MAR-1996;   96WO-US03896.
XX
PR   23-MAR-1995;   95US-0410540.
XX
PA   (REGC ) UNIV CALIFORNIA.
PA   (UYPE-) UNIV PENNSYLVANIA.
XX
PI   Lin D,  Miller WL,  Strauss JF;
XX
DR   WPI; 1996-443130/44.
DR   N-PSDB; AAT62021.
XX
PT   Isolated human steroidogenesis acute regulatory protein gene - used
PT   for detection of mutation(s) of this gene that cause congenital
PT   lipoid adrenal hyperplasia
XX
     Claim 13; Page  ; 89pp; English.

     The present sequence is the human steroidogenesis acute regulatory
CC   protein (hStAR), congenital lipoid adrenal hyperplasia (CAH)
CC   associated mutant, Ala218Val. Analysis of the hStAR gene for such a
CC   mutation is useful in the detection (e.g. prenatally) of CAH
CC   associated genetic defects, or its transmission to children. CAH
CC   can be treated by protein or gene replacement therapy, which can
CC   also be used to prevent or treat hypercholesterolaemia.
XX
SQ   Sequence   285 AA;


  Query Match             100.0%;  Score 38;  DB 17;  Length 285;
  Best Local Similarity   100.0%;  Pred. No. 9.5;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 NQEGWK 6
            ||||||
Db       92 nqeqwk 97


RESULT    9
```

```
AAW13237
ID   AAW13237 standard; Protein; 285 AA.
XX
AC   AAW13237;
XX
DT   21-MAY-1997  (first entry)
XX
DE   Human steroidogenesis acute regulatory protein mutant Leu275Pro.
XX
KW   Human; steroidogenesis; acute regulatory protein; hStAR; analysis;
KW   mutation, detection, prenatal, genetic defect, congenital; protein;
KW   lipoid adrenal hyperplasia; treatment; prevention; gene;
KW   replacement therapy; hypercholesterolaemia; mutant.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 275
FT                   /note= "wild type Leu substituted with Pro"
XX
PN   WO9629338-A1.
XX
PD   26-SEP-1996.
XX
PF   22-MAR-1996;   96WO-US03896.
XX
PR   23-MAR-1995;   95US-0410540.
XX
PA   (REGC ) UNIV CALIFORNIA.
PA   (UYPE-) UNIV PENNSYLVANIA.
XX
PI   Lin D,  Miller WL,  Strauss JF;
XX
DR   WPI; 1996-443130/44.
DR   N-PSDB; AAT62022.
XX
PT   Isolated human steroidogenesis acute regulatory protein gene - used
PT   for detection of mutation(s) of this gene that cause congenital
PT   lipoid adrenal hyperplasia
XX
PS   Claim 13; Page  ; 89pp; English.
XX
CC   The present sequence is the human steroidogenesis acute regulatory
CC   protein (hStAR), congenital lipoid adrenal hyperplasia (CAH)
CC   associated mutant, Leu275Pro. Analysing the hStAR gene for such a
CC   mutation is useful in the detection (e.g. prenatally) of CAH
CC   associated genetic defects, or its transmission to children. CAH
CC   can be treated by protein or gene replacement therapy, which can
CC   also be used to prevent or treat hypercholesterolaemia.
XX
SQ   Sequence   285 AA;


  Query Match             100.0%;  Score 38;  DB 17;  Length 285;
  Best Local Similarity   100.0%;  Pred. No. 9.5;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 NQEGWK 6
            ||||||
Db       92 nqeqwk 97


RESULT   10
AAW00814
ID   AAW00814 standard; Protein; 285 AA
XX
AC   AAW00814;
XX
DT   21-MAY-1997  (first entry)
XX
DE   Human steroidogenesis acute regulatory protein.
XX
KW   Human; steroidogenesis; acute regulatory protein; hStAR; analysis;
KW   mutation; detection; prenatal; genetic defect; [illegible]
KW   lipoid adrenal hyperplasia; treatment; prevention; [illegible]
KW   replacement therapy; hypercholesterolaemia.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Region          [illegible]
FT                   /note= "potential site for protein kinase [illegible]
FT                   mediated phosphorylation"
FT   Region          13
FT                   /note= "potential site for protein kinase [illegible]
FT                   mediated phosphorylation"
FT   Region          [illegible]
FT                   /note= "potential site for protein kinase [illegible]
FT                   mediated phosphorylation"
FT   Region          [illegible]
FT                   /note= "potential site for protein kinase [illegible]
FT                   mediated phosphorylation"
FT   Region          110
FT                   /note= "potential site for protein kinase [illegible]
FT                   mediated phosphorylation"
FT   Region          86
FT                   /note= "potential site for protein kinase [illegible]
FT                   mediated phosphorylation"
FT   Region          [illegible]
FT                   /note= "potential site for protein kinase [illegible]
FT                   mediated phosphorylation"
FT   Region          [illegible]
FT                   /note= "potential site for protein kinase [illegible]
FT                   mediated phosphorylation"
XX
PN   WO96[illegible] A1
XX
PD   26 SEP 1996.
XX
PF   22 MAR 1996;   96WO-[illegible]
XX
PR   23 MAR 1995;   95US-[illegible]
XX
PA   (REGC ) UNIV CALIFORNIA.
PA   (UYPE ) UNIV PENNSYLVANIA
XX
PI   Lin D,  Miller WL,  Strauss JF;
XX
DR   WPI; 1996-443[illegible]/44.
DR   N-PSDB; AAT[illegible]
XX
PT   Isolated human steroidogenesis acute regulatory protein gene - used
PT   for detection of mutation(s) of this gene that cause congenital
PT   lipoid adrenal hyperplasia
XX
PS   Claim 1; Page [illegible]; [illegible]pp; English.
XX
CC   The present sequence is the human steroidogenesis acute
CC   regulatory protein (hStAR). The hStAR gene can be analysed for
CC   mutations to detect (e.g. prenatally) genetic defects associated
CC   with congenital lipoid adrenal hyperplasia (CAH), [illegible]
CC   transmission to children. CAH can be treated by protein or gene
CC   replacement therapy, which can also be used to prevent or treat
CC   hypercholesterolaemia.
CC   A human adrenal cortex cDNA library was screened with [illegible]
CC   probe to isolate a 1.[illegible] kb insert, encoding [illegible]
CC   residue protein. When [illegible]
CC   COS-1 cells cotransfected with pStAR and pADX [illegible]
CC   level of pregnenolone [illegible]
CC   20 alpha hydroxycholesterol.
XX
SQ   Sequence   285 AA;

  Query Match             [illegible]
  Best Local Similarity   [illegible]
```

```
  Matches [illegible]  Conservative [illegible]  Mismatches [illegible]  Indels [illegible]  Gaps [illegible]

[illegible]

[illegible]


RESULT [illegible]
AAW[illegible]
ID   AAW[illegible] standard; Protein; 285 AA.
XX
AC   AAW[illegible]
XX
DT   [illegible] MAY [illegible]
XX
DE   Human steroidogenesis acute regulatory protein mutant [illegible]
XX
KW   Human; steroidogenesis acute regulatory protein; hStAR; analysis;
KW   mutation; detection; prenatal; genetic defect; congenital; protein;
KW   lipoid adrenal hyperplasia; treatment; prevention; gene;
KW   replacement therapy; hypercholesterolaemia; mutant.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference [illegible]
FT                   /note= "wild type [illegible] substituted with [illegible]"
XX
PN   WO[illegible] A1.
XX
PD   [illegible]
XX
PF   [illegible] MAR [illegible]
XX
PR   [illegible] MAR [illegible]
XX
PA   (REGC ) UNIV CALIFORNIA.
PA   (UYPE ) UNIV PENNSYLVANIA.
XX
PI   Lin D,  Miller WL,  Strauss JF;
XX
DR   WPI; 1996-44[illegible]/44.
DR   N-PSDB; AAT[illegible]
XX
PT   Isolated human steroidogenesis acute regulatory protein gene - used
PT   for detection of mutation(s) of this gene that cause congenital
PT   lipoid adrenal hyperplasia
XX
PS   Claim [illegible]; Page [illegible]; [illegible]pp; English.
XX
CC   The present sequence is a human steroidogenesis acute regulatory
CC   protein (hStAR) congenital lipoid adrenal hyperplasia (CAH)
CC   associated mutant, [illegible]. Analysing the hStAR gene for such a
CC   mutation is useful in the detection (e.g. prenatally) of CAH
CC   associated genetic defects, or its transmission to children. CAH
CC   can be treated by protein or gene replacement therapy, which can
CC   also be used to prevent or treat hypercholesterolaemia.
XX
SQ   Sequence   [illegible] AA;


  Query Match             [illegible]
  Best Local Similarity   [illegible]
  Matches [illegible]  Conservative [illegible]  Mismatches [illegible]  Indels [illegible]  Gaps [illegible]

[illegible]

[illegible]


RESULT [illegible]
AAW[illegible]
ID   AAW[illegible] standard; Protein; [illegible] AA.
```

```
XX
AC   AAW13234;
XX
DT   21-MAY-1997  (first entry)
XX
DE   Human steroidogenesis acute regulatory protein mutant Arg182Leu.
XX
KW   Human; steroidogenesis; acute regulatory protein; hStAR; analysis;
KW   mutation; detection; prenatal; genetic defect; congenital; protein;
KW   lipoid adrenal hyperplasia; treatment; prevention; gene;
KW   replacement therapy; hypercholesterolaemia; mutant.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 182
FT                   /note= "wild type Arg substituted with Leu"
XX
PN   WO9629338-A1.
XX
PD   26-SEP-1996.
XX
     22-MAR-1996;   96WO-US03896.

PR   23-MAR-1995;   95US-0410540.
XX
PA   (REGC ) UNIV CALIFORNIA.
PA   (UYPE-) UNIV PENNSYLVANIA.
XX
PI   Lin D,  Miller WL,  Strauss JF;
XX
DR   WPI; 1996-443130/44.
DR   N-PSDB; AAT62019.
XX
PT   Isolated human steroidogenesis acute regulatory protein gene - used
PT   for detection of mutation(s) of this gene that cause congenital
PT   lipoid adrenal hyperplasia
XX
PS   Claim 13; Page -; 89pp; English.
XX
CC   The present sequence is the human steroidogenesis acute regulatory
CC   protein (hStAR). congenital lipoid adrenal hyperplasia (CAH)
CC   associated mutant, Arg182Leu. Analysing the hStAR gene for such a
CC   mutation is useful in the detection (e.g. prenatally) of CAH
CC   associated genetic defects, or its transmission to children. CAH
CC   can be treated by protein or gene replacement therapy, which can
CC   also be used to prevent or treat hypercholesterolaemia.
XX
SQ   Sequence   285 AA;

  ery Match             100.0%;  Score 38;  DB 17;  Length 285;
  est Local Similarity   100.0%;  Pred. No. 9.5;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 NQEGWK 6
           ||||||
Db      92 nqegwk 97

RESULT  13
AAW13235
ID   AAW13235 standard; Protein; 285 AA.
XX
AC   AAW13235;
XX
DT   21-MAY-1997  (first entry)
XX
DE   Human steroidogenesis acute regulatory protein mutant Glu169Lys.
XX
KW   Human; steroidogenesis; acute regulatory protein; hStAR; analysis;
KW   mutation; detection; prenatal; genetic defect; congenital; protein;
KW   lipoid adrenal hyperplasia; treatment; prevention; gene;
KW   replacement therapy; hypercholesterolaemia; mutant.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 169
FT                   /note= "wild type Glu substituted with Lys"
XX
PN   WO9629338-A1.
XX
PD   26-SEP-1996.
XX
PF   22-MAR-1996;   96WO-US03896.
XX
PR   23-MAR-1995;   95US-0410540.
XX
PA   (REGC ) UNIV CALIFORNIA.
PA   (UYPE-) UNIV PENNSYLVANIA.
XX
PI   Lin D,  Miller WL,  Strauss JF;
XX
DR   WPI; 1996-443130/44.
DR   N-PSDB; AAT62020.
XX
PT   Isolated human steroidogenesis acute regulatory protein gene - used
PT   for detection of mutation(s) of this gene that cause congenital
PT   lipoid adrenal hyperplasia
XX
PS   Claim 13; Page -; 89pp; English.
XX
CC   The present sequence is the human steroidogenesis acute regulatory
CC   protein (hStAR), congenital lipoid adrenal hyperplasia (CAH)
CC   associated mutant, Glu169Lys. Analysing the hStAR gene for such a
CC   mutation is useful in the detection (e.g. prenatally) of CAH
CC   associated genetic defects, or its transmission to children. CAH
CC   can be treated by protein or gene replacement therapy, which can
CC   also be used to prevent or treat hypercholesterolaemia.
XX
SQ   Sequence   285 AA;

  Query Match             100.0%;  Score 38;  DB 17;  Length 285;
  Best Local Similarity   100.0%;  Pred. No. 9.5;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 NQEGWK 6
           ||||||
Db      92 nqegwk 97

RESULT  14
AAY72838
ID   AAY72838 standard; Protein; 285 AA.
XX
AC   AAY72838;
XX
DT   31-MAY-2001  (first entry)
XX
DE   Human steroidogenic acute regulatory (StAR) protein.
XX
KW   Human; steroidogenic acute regulatory protein; StAR;
KW   gene therapy; steroid hormone-dependent disorder; LCAH;
KW   lipoid congenital adrenal hyperplasia; sexual maturation;
KW   infertility; androgen-responsive tumour; precocious puberty;
KW   hypogonadotropic hypogonadism; McCune-Albright syndrome;
KW   adrenal-hypoplasia congenita; cytostatic.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 156
FT                   /note= "Encoded by "A" this feature corresponds
FT                   to the translation exception to the sequence shown
FT                              [illegible] (PKA [illegible])"
FT   Modified site              [illegible]
FT                              /note= "PKA phosphorylation site"
FT   Modified site              [illegible]
FT                              /note= "PK phosphorylation site"
FT   Modified site              [illegible]
FT                              /note= "PKA phosphorylation site"
FT   Modified site              [illegible]
FT                              /note= "P34cdc2 kinase (?) phosphorylation site"
FT   Modified site              [illegible]8
FT                              /note= "PKC phosphorylation site"
FT   Modified site              [illegible]
FT                              /note= "PKA phosphorylation site"
XX
PN   US6194555 B1
XX
PD   27 FEB 2001.
XX
PF   07 JUN 1996;   96US-[illegible].
XX
PR   04 NOV 1995;   95US-[illegible].
XX
PA   (UYTE-) UNIV TEXAS TECH HEALTH SCI CENT
XX
PI   Stocco DM, Clark BJ
XX
DR   WPI; 2001-217[illegible].
DR   N-PSDB; AAD[illegible], AAD[illegible].
XX
PT   Novel nucleic acid encoding steroidogenic acute regulatory protein,
PT   useful for treating steroid hormone dependent disorders
XX
PS   Example 14; Fig [illegible]; 41pp; English.
XX
CC   The invention relates to steroidogenic acute regulatory protein
CC   (StAR) and nucleic acid molecules encoding them. StAR proteins
CC   are used for regulating steroidogenesis. More particularly, these
CC   proteins are involved in regulating cholesterol transport into
CC   the mitochondria of a cell. The nucleic acid molecules of the
CC   invention are useful in gene therapy for the treatment of steroid
CC   hormone dependent disorders, particularly congenital lipoid
CC   adrenal hyperplasia (CLAH), infertility, sexual maturation,
CC   androgen responsive tumours, precocious puberty, McCune-Albright
CC   syndrome, adrenal hypoplasia congenita, [illegible]
CC   hypogonadism. They are also useful for recombinant production
CC   of the steroidogenic acute regulatory protein.
CC   The present sequence is human steroidogenic acute regulatory
CC   protein (StAR)
XX
SQ   Sequence  28[illegible] AA;

  Query Match              [illegible]
  Best Local Similarity    [illegible]
  Matches  [illegible]

Qy          1 NQEGWK 6
              [illegible]
Db         92 nqegwk 97


RESULT 15
AAY28814
ID   AAY28814 standard; Protein; [illegible] AA.
XX
AC   AAY28814;
XX
DT   17-JAN-2000  (first entry)
XX
DE   pk855_1 secreted protein.
XX
KW   Clone pk855_1; pk855_1 [illegible]
KW   secreted protein; transmembrane domain; [illegible]
KW   [illegible] signal peptide; hydrophobic [illegible]
KW   nutritional activity; cell proliferation; immune stimulation;
KW   [illegible]
XX
OS   Homo sapiens.
XX
FH   Key                Location/Qualifiers
FT   Peptide            [illegible]
FT                      /note= "Predicted signal peptide"
FT   Protein            [illegible]
FT                      /note= "Mature pk855_1 protein"
XX
PN   WO[illegible]-A1.
XX
PD   [illegible]
XX
PF   [illegible]-MAR-[illegible]   [illegible]
XX
PR   [illegible]-MAR-[illegible]   [illegible]
PR   [illegible]-MAR-[illegible]   [illegible]
XX
PA   (GEMY-) GENETICS INST INC
XX
PI   [illegible] Merberg D, Treacy M, Agostino MJ, [illegible] BJ;
XX
DR   WPI; [illegible]
DR   N-PSDB; AAX00854
XX
PT   Polynucleotides encoding secreted human proteins, derived from human
PT   adult brain, human foetal brain, human foetal kidney and human adult
PT   [illegible] cDNA libraries
XX
PS   [illegible]; English.
XX
CC   The present sequence is the pk855_1 secreted protein encoded by the cDNA
CC   clone pk855_1. pk855_1 was isolated from a human foetal kidney cDNA
CC   library using methods specific for secreted protein cDNAs. The leader
CC   sequence or signal peptide acts as a transmembrane domain due to its
CC   hydrophobic nature. The TopPred II computer program predicts an
CC   additional potential transmembrane domain centred around amino acid [illegible].
CC   The polynucleotide and protein may effect nutritional activity, cytokine
CC   and cell proliferation, immune stimulation or suppression, tumour [illegible]
CC   [illegible] with tumour inhibition [illegible]
XX
SQ   Sequence  [illegible] AA;

  Query Match              [illegible]  Length [illegible]
  Best Local Similarity    [illegible]
  Matches  [illegible]  Mismatches  [illegible]  Indels  [illegible]  Gaps  [illegible]

Qy          [illegible] NQEGWK [illegible]
              [illegible]
Db         [illegible] nqegwk [illegible]


Search completed: [illegible]
[illegible]
```

```
                     GenCore version 4.5
              Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on:          August 22, 2001, 14:30:11 ; Search time 55.29 Seconds
                                             (without alignments)
                                             2.234 Million cell updates/sec

Title:           US-08-659-254-6
Perfect score:   38
Sequence:        1 NQEGWK 6

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        197339 seqs, 20590346 residues

Total number of hits satisfying chosen parameters:      197339

 mum DB seq length: 0
 mum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :        Issued_Patents_AA:*
                1:  /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
                2:  /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
                3:  /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
                4:  /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
                5:  /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
                6:  /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                 SUMMARIES

                 %
Result         Query
  No.   Score  Match Length DB  ID                          Description
---------------------------------------------------------------------------
    1      38  100.0      6  2  US-08-538-960-6             Sequence 6, Appli
    2      38  100.0      6  4  US-08-659-254-6             Sequence 6, Appli
    3      38  100.0     11  2  US-08-538-960-8             Sequence 8, Appli
    4      38  100.0     11  4  US-08-659-254-8             Sequence 8, Appli
    5      38  100.0     12  2  US-08-538-960-4             Sequence 4, Appli
    6      38  100.0     12  4  US-08-659-254-4             Sequence 4, Appli
    7      38  100.0    276  4  US-08-659-254-2             Sequence 2, Appli
    8      38  100.0    284  2  US-08-538-960-2             Sequence 2, Appli
    9      38  100.0    285  1  US-08-41[illegible]-540-2    Sequence 2, Appli
   10      38  100.0    285  4  US-08-659-254-18            Sequence 18, Appl
   11      33   86.8    514  1  US-08-36[illegible]-920-21   Sequence 21, Appl
   12      33   86.8    514  1  US-08-47[illegible]-939-21   Sequence 21, App.
   13      33   86.8    514  1  US-08-48[illegible]-432-21   Sequence 21, App.
   14      32   84.2     96  2  US-08-48[illegible]-537-12   Sequence 12, App.
   15      32   84.2     96  2  US-08-47[illegible]-15-12    Sequence 12, App.
   16      32   84.2     96  5  PCT-US9[illegible]-552-12    Sequence 12, Appl
   17      32   84.2    333  1  US-08-208-[illegible]07A-9   Sequence 9, Appli
   18      32   84.2    333  1  US-08-339-121B-5            Sequence 5, Appli
   19      32   84.2    333  3  US-08-883-[illegible]26-3    Sequence 3, Appli
   20      32   84.2    333  3  US-09-120-465-65            Sequence 65, Appl
   21      32   84.2    333  4  US-09-51[illegible]-29-65    Sequence 65, Appl
   22      32   84.2    333  5  PCT-US95-[illegible]1-820-5  Sequence 5, Appli
   23      32   84.2    334  3  US-08-88[illegible]-526-3    Sequence 3, Appli
   24      32   84.2    334  3  US-08-88[illegible]-526-4    Sequence 4, Appli
   25      32   84.2   2205  1  US-08-09[illegible]-453B-2   Sequence 2, Appli
   26      30   78.9    199  2  US-08-924-759-6             Sequence 6, Appli
   27      30   78.9    199  3  US-09-248-[illegible]35-6    Sequence 6, Appli
   28      30   78.9    415  5  PCT-US93-[illegible]0601-2   Sequence 2, Appli
   29      30   78.9    415  5  PCT-US94-[illegible]107A-2   Sequence 2, Appli
   30      30   78.9    426  5  PCT-US94-07107A-7           Sequence 7, Appli
   31      30   78.9    452  4  US-09-[illegible]-35         Sequence 35, Appl
   32      29   76.3     20  2  US-08-[illegible]-366-13     Sequence 13, Appl
   33      29   76.3    108  2  US-08-[illegible]-87-37      Sequence 37, Appl
   34      29   76.3    108  4  US-08-6[illegible]-37        Sequence 37, Appl
   35      29   76.3   1429  1  US-07-64[illegible]-2       Sequence 2, Appli
   36      29   76.3   1429  2  US-08-[illegible]-486A-13    Sequence 13, Appl
   37      29   76.3   1429  2  US-08-[illegible]-11        Sequence 11, Appl
   38      29   76.3   1429  4  US-08-880-342-13            Sequence 13, Appl
   39      29   76.3   1430  2  US-08-705-625-4             Sequence 4, Appli
   40      29   76.3   1430  3  US-09-[illegible]-998-5     Sequence 5, Appli
   41      29   76.3   1430  4  US-09-[illegible]-574-4     Sequence 4, Appli
   42      29   76.3   1433  2  US-08-66[illegible]-486A-21  Sequence 21, Appl
   43      29   76.3   1433  4  US-09-12[illegible]-708-4   Sequence 4, Appli
   44      29   76.3   1433  4  US-09-12[illegible]-624-4   Sequence 4, Appli
   45      29   76.3   1433  4  US-08-3[illegible]-342-21   Sequence 21, Appl

                                 ALIGNMENTS

RESULT 1
US-08-538-960-6
; Sequence 6, Application US/08538960
; Patent No. 5872236
;  GENERAL INFORMATION:
;     APPLICANT:  Stocco, Douglas M.
;     APPLICANT:  Clark, Barbara J.
;     TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR
;     TITLE OF INVENTION:  REGULATION OF [illegible]ENESIS
;     NUMBER OF SEQUENCES:  14
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Akin, Gump, Strauss, Hauer & Feld, L.L.P.
;       STREET:  [illegible] First Bank Place [illegible] Avenue
;       CITY:  Austin
;       STATE:  TX
;       COUNTRY:  U.S.A.
;       ZIP:  78701
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/538,960
;       FILING DATE:  Concurrently Herewith
;       CLASSIFICATION:  530
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Mayfield, Denise L.
;       REGISTRATION NUMBER:  33,732
;       REFERENCE/DOCKET NUMBER:  [illegible]/DLM
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  512/499-6200
;       TELEFAX:  512/499-6290
;   INFORMATION FOR SEQ ID NO:  6:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  6 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
US-08-538-960-6

  Query Match             100.0%;  Score 38;  DB 2;  Length 6;
  Best Local Similarity   100.0%;  Pred. No. 1.5e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 NQEGWK 6
              ||||||
Db          1 NQEGWK 6
```

```
RESULT    2
US-08-659-254-6
; Sequence [illegible], Application [illegible]
; Patent No. 61[illegible]
;  GENERAL INFORMATION:
;     APPLICANT: [illegible] M
;     APPLICANT: Clark, [illegible] Barbara [illegible]
;     TITLE OF INVENTION: COMPOSITIONS AND [illegible]
;     TITLE OF INVENTION: OF STEROIDOGENESIS
;     NUMBER OF SEQUENCES: 1[illegible]
;     CORRESPONDENCE ADDRESS:
;        ADDRESSEE: Akin, Gump, Strauss, Hauer & Feld [illegible]
;        STREET: 1900 Frost Bank Plaza, [illegible]
;        CITY: Austin
;        STATE: TX
;        COUNTRY: U.S.A.
;        ZIP: 78701
;     COMPUTER READABLE FORM:
;        MEDIUM TYPE: Floppy disk
;        COMPUTER: IBM PC compatible
;        OPERATING SYSTEM: [illegible]/MS-DOS
;        SOFTWARE: PatentIn Release #1.0, Version #[illegible]
;     CURRENT APPLICATION DATA:
;        APPLICATION NUMBER: US/08/[illegible]
;        FILING DATE: [illegible]
;        CLASSIFICATION: [illegible]
;     PRIOR APPLICATION DATA:
;        APPLICATION NUMBER: US [illegible]
;        FILING DATE: [illegible]
;     ATTORNEY/AGENT INFORMATION:
;        NAME: Mayfield, Denise L.
;        REGISTRATION NUMBER: 33,732
;        REFERENCE/DOCKET NUMBER: [illegible]
;     TELECOMMUNICATION INFORMATION:
;        TELEPHONE: 512/499-6200
;        TELEFAX: 512/[illegible]
;  INFORMATION FOR SEQ ID NO: [illegible]
;     SEQUENCE CHARACTERISTICS:
;        LENGTH: [illegible] amino acids
;        TYPE: amino acid
;        STRANDEDNESS: single
;        TOPOLOGY: linear
;     MOLECULE TYPE: peptide
US-08-659-254-6

  Query Match             [illegible]  Score [illegible]
  Best Local Similarity   [illegible]  Pred. No. [illegible]
  Matches [illegible]  Conservative [illegible]  Mismatches [illegible]  Indels [illegible]  Gaps [illegible]

Qy          1 NQFGWK [illegible]
              ||||||
Db          1 NQFGWK [illegible]


RESULT    3
US-08-538-966-8
; Sequence 8, Application [illegible]
; Patent No. 5872[illegible]
;  GENERAL INFORMATION:
;     APPLICANT: Stocco, Douglas M
;     APPLICANT: Clark, Barbara [illegible]
;     TITLE OF INVENTION: COMPOSITIONS AND METHODS [illegible]
;     TITLE OF INVENTION: REGULATION OF STEROIDOGENESIS
;     NUMBER OF SEQUENCES: [illegible]
;     CORRESPONDENCE ADDRESS:
;        ADDRESSEE: Akin, Gump, Strauss, Hauer & Feld [illegible]
;        STREET: 1900 Frost Bank Plaza, [illegible]
;        CITY: Austin
;        STATE: TX
;        COUNTRY: U.S.A.
;        ZIP: [illegible]
;     COMPUTER READABLE FORM:
;        MEDIUM TYPE: Floppy disk
;        COMPUTER: IBM PC compatible
;        OPERATING SYSTEM: PC-DOS/MS-DOS
;        SOFTWARE: PatentIn Release #[illegible], Version #[illegible]
;     CURRENT APPLICATION DATA:
;        APPLICATION NUMBER: US/[illegible]
;        FILING DATE: [illegible]
;        CLASSIFICATION: [illegible]
;     ATTORNEY/AGENT INFORMATION:
;        NAME: Mayfield, Denise [illegible]
;        REGISTRATION NUMBER: [illegible]
;        REFERENCE/DOCKET NUMBER: [illegible]
;     TELECOMMUNICATION INFORMATION:
;        TELEPHONE: [illegible]
;        TELEFAX: 512/[illegible]
;  INFORMATION FOR SEQ ID NO: [illegible]
;     SEQUENCE CHARACTERISTICS:
;        LENGTH: [illegible] amino acids
;        TYPE: amino acid
;        STRANDEDNESS: single
;        TOPOLOGY: linear
;     MOLECULE TYPE: peptide
US-[illegible]

  Query Match             [illegible]  Score [illegible]  Length [illegible]
  Best Local Similarity   [illegible]  Pred. No. [illegible]
  Matches [illegible]  Conservative [illegible]  Mismatches [illegible]  Indels [illegible]  Gaps [illegible]

Qy          1 NQFGWK [illegible]

Db          4 NQFGWK [illegible]


RESULT    4
US-[illegible]
; Sequence [illegible], Application [illegible]
; Patent No. [illegible]
;  GENERAL INFORMATION:
;     APPLICANT: Stocco, Douglas M
;     APPLICANT: Clark, [illegible] Barbara [illegible]
;     TITLE OF INVENTION: COMPOSITIONS AND METHODS [illegible] REGULATION
;     TITLE OF INVENTION: OF STEROIDOGENESIS
;     NUMBER OF SEQUENCES: [illegible]
;     CORRESPONDENCE ADDRESS:
;        ADDRESSEE: Akin, Gump, Strauss, Hauer & Feld, L.L.P.
;        STREET: 1900 Frost Bank Plaza, 816 Congress Avenue
;        CITY: Austin
;        STATE: TX
;        COUNTRY: U.S.A.
;        ZIP: [illegible]
;     COMPUTER READABLE FORM:
;        MEDIUM TYPE: Floppy disk
;        COMPUTER: IBM PC compatible
;        OPERATING SYSTEM: PC-DOS/MS-DOS
;        SOFTWARE: PatentIn Release #1.0, Version #[illegible]
;     CURRENT APPLICATION DATA:
;        APPLICATION NUMBER: US/08/[illegible]254
;        FILING DATE: [illegible]
;        CLASSIFICATION: [illegible]
;     PRIOR APPLICATION DATA:
;        APPLICATION NUMBER: [illegible]
;        FILING DATE: [illegible]
;     ATTORNEY/AGENT INFORMATION:
;        NAME: Mayfield, Denise L.
;        REGISTRATION NUMBER: [illegible]
;        REFERENCE/DOCKET NUMBER: [illegible]
;     TELECOMMUNICATION INFORMATION:
;        TELEPHONE: 512/[illegible]
;        TELEFAX: [illegible]
;  INFORMATION FOR SEQ ID NO: [illegible]
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  11 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
US-08-659-254-8


  Query Match             100.0%;  Score 38;  DB 4;  Length 11;
  Best Local Similarity   100.0%;  Pred. No. 0.21;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 NQEGWK 6
              ||||||
Db          4 NQEGWK 9


RESULT    5
US-08-538-960-4
; Sequence 4, Application US/08538960
; Patent No. 5872230
;  GENERAL INFORMATION:
;     APPLICANT:  Stocco, Douglas M.
;     APPLICANT:  Clark, Barbara J.
;     TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR
;     TITLE OF INVENTION:  REGULATION OF STEROIDOGENESIS
;     NUMBER OF SEQUENCES:  14
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Akin, Gump, Strauss, Hauer & Feld, L.L.P.
;       STREET:  1900 Frost Bank Plaza, 816 Congress Avenue
;       CITY:  Austin
;       STATE:  TX
;       COUNTRY:  U.S.A.
;       ZIP:  78701
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/538,960
;       FILING DATE:  Concurrently Herewith
;       CLASSIFICATION:  530
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Mayfield, Denise L.
;       REGISTRATION NUMBER:  33,732
;       REFERENCE/DOCKET NUMBER:  43375.0002/DLM
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  512/499-6200
;       TELEFAX:  512/499-6290
;  INFORMATION FOR SEQ ID NO:  4:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  12 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
US-08-538-960-4


  Query Match             100.0%;  Score 38;  DB 2;  Length 12;
  Best Local Similarity   100.0%;  Pred. No. 0.23;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 NQEGWK 6
              ||||||
Db          7 NQEGWK 12


RESULT    6
US-08-659-254-4
; Sequence 4, Application US/08659254
; Patent No. 6194555
;  GENERAL INFORMATION:
;     APPLICANT:  Stocco, Douglas M.
;     APPLICANT:  Clark, Dr. Barbara J.
;     TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR REGULATION
;     TITLE OF INVENTION:  OF STEROIDOGENESIS
;     NUMBER OF SEQUENCES:  19
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Akin, Gump, Strauss, Hauer & Feld, L.L.P.
;       STREET:  1900 Frost Bank Plaza, 816 Congress Avenue
;       CITY:  Austin
;       STATE:  TX
;       COUNTRY:  U.S.A.
;       ZIP:  78701
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/659,254
;       FILING DATE:  07-JUN-1996
;       CLASSIFICATION:  530
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/538,960
;       FILING DATE:  04-OCT-1995
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Mayfield, Denise L.
;       REGISTRATION NUMBER:  33,732
;       REFERENCE/DOCKET NUMBER:  43375.0006
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  512/499-6200
;       TELEFAX:  512/499-6290
;  INFORMATION FOR SEQ ID NO:  4:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  12 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
US-08-659-254-4


  Query Match             100.0%;  Score 38;  DB 4;  Length 12;
  Best Local Similarity   100.0%;  Pred. No. 0.23;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 NQEGWK 6
              ||||||
Db          7 NQEGWK 12


RESULT    7
US-08-659-254-2
; Sequence 2, Application US/08659254
; Patent No. 6194555
;  GENERAL INFORMATION:
;     APPLICANT:  Stocco, Douglas M.
;     APPLICANT:  Clark, Dr. Barbara J
;     TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR REGULATION
;     TITLE OF INVENTION:  OF STEROIDOGENESIS
;     NUMBER OF SEQUENCES:  19
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Akin, Gump, Strauss, Hauer & Feld, L.L.P.
;       STREET:  1900 Frost Bank Plaza, 816 Congress Avenue
;       CITY:  Austin
;       STATE:  TX
;       COUNTRY:  U.S.A.
;       ZIP:  78701
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;          SOFTWARE:  PatentIn Release #[illegible] Version #[illegible]
;       CURRENT APPLICATION DATA:
;          APPLICATION NUMBER: US/08/[illegible]
;          FILING DATE: [illegible]
;          CLASSIFICATION: [illegible]
;       PRIOR APPLICATION DATA:
;          APPLICATION NUMBER: US 08/[illegible]
;          FILING DATE: [illegible]
;       ATTORNEY/AGENT INFORMATION:
;          NAME:  Mayfield, Denise L
;          REGISTRATION NUMBER: [illegible]
;          REFERENCE/DOCKET NUMBER: [illegible]
;       TELECOMMUNICATION INFORMATION:
;          TELEPHONE: [illegible]/49[illegible]
;          TELEFAX: [illegible]
;    INFORMATION FOR SEQ ID NO: 2:
;       SEQUENCE CHARACTERISTICS:
;          LENGTH: 2[illegible] amino acids
;          TYPE: amino acid
;          STRANDEDNESS: single
;          TOPOLOGY: linear
;       MOLECULE TYPE: protein
US-08-659-254-2

  Query Match             [illegible]%;  Score [illegible];  DB [illegible];  Length [illegible];
  Best Local Similarity   100.0%;  Pred. No. [illegible];
  Matches   [illegible];  Conservative   [illegible];  Mismatches   [illegible];  Indels   [illegible];  Gaps   [illegible];

Qy          1 NQEGWK 6
              ||||||
Db         41 NQEGWK 46


RESULT 8
US-08-538-960-2
; Sequence 2, Application US/08538960
; Patent No. 5872230
;  GENERAL INFORMATION:
;       APPLICANT:  Stocco, Douglas M
;       APPLICANT:  Clark, Barbara J
;       TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR
;       TITLE OF INVENTION:  REGULATION OF STEROIDOGENESIS
;       NUMBER OF SEQUENCES:  14
;       CORRESPONDENCE ADDRESS:
;          ADDRESSEE:  Akin, Gump, Strauss, Hauer & Feld, [illegible]
;          STREET:  1900 Frost Bank Plaza, 816 Congress Avenue
;          CITY:  Austin
;          STATE:  TX
;          COUNTRY:  U.S.A.
;          ZIP:  78701
;       COMPUTER READABLE FORM:
;          MEDIUM TYPE:  Floppy disk
;          COMPUTER:  IBM PC compatible
;          OPERATING SYSTEM:  PC-DOS/MS-DOS
;          SOFTWARE:  PatentIn Release #[illegible] Version #[illegible]
;       CURRENT APPLICATION DATA:
;          APPLICATION NUMBER:  US/08/[illegible]
;          FILING DATE:  Concurrently herewith
;          CLASSIFICATION:  [illegible]
;       ATTORNEY/AGENT INFORMATION:
;          NAME:  Mayfield, Denise L
;          REGISTRATION NUMBER:  [illegible]
;          REFERENCE/DOCKET NUMBER:  [illegible]
;       TELECOMMUNICATION INFORMATION:
;          TELEPHONE:  [illegible] 499-[illegible]
;          TELEFAX:  512/4[illegible]
;    INFORMATION FOR SEQ ID NO: [illegible]
;       SEQUENCE CHARACTERISTICS:
;          LENGTH:  284 amino acids
;          TYPE:  amino acid
;          STRANDEDNESS:  [illegible]
;          TOPOLOGY:  linear
```

```
;       MOLECULE TYPE: [illegible]
US-[illegible]

  Query Match             [illegible]
  Best Local Similarity   [illegible];  Pred. No. [illegible];
  Matches   [illegible];  Conservative   [illegible];  Mismatches   [illegible];  Indels   [illegible];  Gaps   [illegible];

Qy          [illegible] NQEGWK [illegible]

Db          [illegible] NQEGWK [illegible]


RESULT 9
US-[illegible]
; Sequence [illegible], Application US/[illegible]
; Patent No. [illegible]
;  GENERAL INFORMATION:
;       APPLICANT:  Miller, Walter L
;       APPLICANT:  [illegible]
;       APPLICANT:  Strauss III, Jerome F.
;       TITLE OF INVENTION:  IDENTIFICATION OF GENE MUTATIONS
;       TITLE OF INVENTION:  ASSOCIATED WITH CONGENITAL LIPOID ADRENAL HYPERPLASIA
;       NUMBER OF SEQUENCES:  [illegible]
;       CORRESPONDENCE ADDRESS:
;          ADDRESSEE:  [illegible] Edward Castro Huddleson & Tatum
;          STREET:  [illegible]
;          CITY:  [illegible] Alto
;          STATE:  CA
;          COUNTRY:  US
;          ZIP:  [illegible]
;       COMPUTER READABLE FORM:
;          MEDIUM TYPE:  Floppy disk
;          COMPUTER:  IBM PC compatible
;          OPERATING SYSTEM:  PC-DOS/MS-DOS
;          SOFTWARE:  PatentIn Release #1.[illegible], Version #1.[illegible]
;       CURRENT APPLICATION DATA:
;          APPLICATION NUMBER:  US/0[illegible]4
;          FILING DATE:  [illegible]-MAR-[illegible]
;          CLASSIFICATION:  [illegible]
;       ATTORNEY/AGENT INFORMATION:
;          NAME:  Neeley, [illegible]
;          REGISTRATION NUMBER:  [illegible]
;          REFERENCE/DOCKET NUMBER:  [illegible]
;       TELECOMMUNICATION INFORMATION:
;          TELEPHONE:  [illegible]
;          TELEFAX:  41[illegible]
;          TELEX:  [illegible]
;    INFORMATION FOR SEQ ID NO: [illegible]
;       SEQUENCE CHARACTERISTICS:
;          LENGTH:  [illegible] amino acids
;          TYPE:  amino acid
;          [illegible]
;       MOLECULE TYPE:  protein
US-[illegible]

  Query Match             [illegible];  Length [illegible];
  Best Local Similarity   [illegible];  Pred. No. [illegible];
  Matches   [illegible];  Conservative   [illegible];  Mismatches   [illegible];  Indels   [illegible];  Gaps   [illegible];

Qy          [illegible] NQEGWK [illegible]
              [illegible]
Db          [illegible] NQEGWK [illegible]


RESULT 1[illegible]
US-[illegible]
; Sequence [illegible], Application US/[illegible]4
; Patent No. [illegible]
;  GENERAL INFORMATION:
;       APPLICANT:  Stocco, Douglas M
;       APPLICANT:  Clark, Barbara J
;    TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR REGULATION
;    TITLE OF INVENTION:  OF STEROIDOGENESIS
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Akin, Gump, Strauss, Hauer & Feld, L.L.P.
;      STREET:  1900 Frost Bank Plaza, 816 Congress Avenue
;      CITY:  Austin
;      STATE:  TX
;      COUNTRY:  U.S.A.
;      ZIP:  78701
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/659,254
;      FILING DATE:  07-JUN-1996
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/538,960
;      FILING DATE:  04-OCT-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Mayfield, Denise L.
;      REGISTRATION NUMBER:  33,732
;      REFERENCE/DOCKET NUMBER:  43375 0000
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  512/499-6200
;      TELEFAX:  512/499-6290
;  INFORMATION FOR SEQ ID NO:  18:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  285 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
US-08-659-254-18


  Query Match              100.0%;  Score 38;  DB 4;  Length 285;
  Best Local Similarity    100.0%;  Pred. No. 5.5;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 NQEGWK 6
           ||||||
Db      92 NQEGWK 97


RESULT  11
US-08-361-920-21
; Sequence 21, Application US/08361920
; Patent No. 5457046
;  GENERAL INFORMATION:
;    APPLICANT:  Woeldike, Helle F.
;    APPLICANT:  Hagen, Frederick
;    APPLICANT:  Hjort, Carsten M.
;    APPLICANT:  Sven, Hastrup
;    TITLE OF INVENTION:  An Enzyme Capable of Degrading Cellulose
;    TITLE OF INVENTION:  or Hemicellulose
;    NUMBER OF SEQUENCES:  85
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 5457046o No. [illegible]
;      STREET:  405 Lexington Avenue, 62nd Floor
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  United States of America
;      ZIP:  10174-6201
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/361,920
;      FILING DATE:
```

```
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/940,8[illegible]
;      FILING DATE:  28-OCT-1992
;      APPLICATION NUMBER:  DK 1159/90
;      FILING DATE:  09-MAY-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/DK91/00124
;      FILING DATE:  08-MAY-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Lambiris, Elias J.
;      REGISTRATION NUMBER:  33,728
;      REFERENCE/DOCKET NUMBER:  3435.204-US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212 867-0123
;      TELEFAX:  212 867-0298
;  INFORMATION FOR SEQ ID NO:  21:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  514 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-361-920-21


  Query Match              86.8%;  Score 33;  DB 1;  Length 514;
  Best Local Similarity    83.3%;  Pred. No. 8[illegible];
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 NQEGWK 6
           | ||||
Db     205 NSEGWK 210


RESULT  12
US-08-479-939-21
; Sequence 21, Application US/08479939
; Patent No. 5686593
;  GENERAL INFORMATION:
;    APPLICANT:  Woeldike, Helle F.
;    APPLICANT:  Hagen, Frederick
;    APPLICANT:  Hjort, Carsten M.
;    APPLICANT:  Sven, Hastrup
;    TITLE OF INVENTION:  An Enzyme Capable of Degrading Cellulose
;    TITLE OF INVENTION:  or Hemicellulose
;    NUMBER OF SEQUENCES:  85
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 5686593o No. 5686593d.ek of No. 5686593th America, Inc.
;      STREET:  405 Lexington Avenue, 62nd Floor
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  United States of America
;      ZIP:  10174-6201
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/479,939
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/361,920
;      FILING DATE:  22-DEC-1994
;      APPLICATION NUMBER:  US 07/940,85[illegible]
;      FILING DATE:  28-OCT-1992
;      APPLICATION NUMBER:  DK 1159/90
;      FILING DATE:  09-MAY-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/DK91/00124
;      FILING DATE:  08 MAY 1991
;    ATTORNEY/AGENT INFORMATION:
```

```
;       NAME: Lambiris, [illegible]
;       REGISTRATION NUMBER: [illegible]
;       REFERENCE/DOCKET NUMBER: [illegible]
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE: [illegible] 867 [illegible]
;       TELEFAX: 21[illegible] 867 [illegible]
;  INFORMATION FOR SEQ ID NO: 21:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH: 511 amino acids
;       TYPE: amino acid
;       TOPOLOGY: linear
;     MOLECULE TYPE: protein
US-08-479-599-21

  Query Match            96.8%; Score [illegible]; DB [illegible]; Length [illegible];
  Best Local Similarity  [illegible]; Pred. No. [illegible];
  Matches [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible];

Qy          1 NQEGWK 6
                |||
Db        205 NQEGWK 210


RESULT 13
US-08-483-132-21
; Sequence 21, Application US/08483132
; Patent No. 5763254
;  GENERAL INFORMATION:
;     APPLICANT: Woeldike, Helle F
;     APPLICANT: Hansen, Frederik
;     APPLICANT: [illegible], Christen M
;     APPLICANT: [illegible]
;     TITLE OF INVENTION: An Enzyme [illegible]
;     TITLE OF INVENTION: [illegible]
;     NUMBER OF SEQUENCES: [illegible]
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE: [illegible]
;       STREET: 4[illegible] Lexington Avenue, [illegible]
;       CITY: New York
;       STATE: New York
;       COUNTRY: United States of America
;       ZIP: 10174-[illegible]
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE: Floppy disk
;       COMPUTER: IBM PC compatible
;       OPERATING SYSTEM: PC-DOS/MS-DOS
;       SOFTWARE: PatentIn Release #1.0, Version #[illegible]
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER: US/08/483,132
;       FILING DATE: [illegible] JUN 199[illegible]
;       CLASSIFICATION: 4[illegible]
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER: US/08/361,[illegible]
;       FILING DATE:
;       APPLICATION NUMBER: US 07/910,[illegible]
;       FILING DATE: [illegible]
;       APPLICATION NUMBER: DK 1158/90
;       FILING DATE: [illegible] MAY [illegible]
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER: [illegible]
;       FILING DATE: [illegible] MAY 1991
;     ATTORNEY/AGENT INFORMATION:
;       NAME: Lambiris, El[illegible]
;       REGISTRATION NUMBER: [illegible]
;       REFERENCE/DOCKET NUMBER: [illegible]
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE: 21[illegible] 867 [illegible]
;       TELEFAX: 212 [illegible]
;  INFORMATION FOR SEQ ID NO: [illegible]:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH: 511 amino acids
;       TYPE: amino acid
;       TOPOLOGY: linear
;     MOLECULE TYPE: protein
US-08-483-132-21

  Query Match            [illegible]; Score [illegible]; DB [illegible]; Length 514;
  Best Local Similarity  [illegible]; Pred. No. [illegible];
  Matches [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible];

Qy          1 NQEGWK 6
                [illegible]
Db        [illegible] NQEGWK [illegible]


RESULT 14
US-[illegible]
; Sequence [illegible], Application US/[illegible]
; Patent No. [illegible]
;  GENERAL INFORMATION:
;     APPLICANT: M[illegible]
;     APPLICANT: Brow[illegible]
;     APPLICANT: [illegible]
;     APPLICANT: [illegible]
;     TITLE OF INVENTION: [illegible]
;     TITLE OF INVENTION: [illegible] Protease Inhibitor
;     NUMBER OF SEQUENCES: [illegible]
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE: [illegible]
;       STREET: [illegible] Wacker [illegible]
;       CITY: [illegible]
;       STATE: [illegible]
;       COUNTRY: USA
;       ZIP: [illegible]
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE: Floppy disk
;       COMPUTER: IBM PC compatible
;       OPERATING SYSTEM: PC-DOS/MS-DOS
;       SOFTWARE: PatentIn Release #1.0, Version #[illegible]
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER: [illegible]
;       FILING DATE:
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER: [illegible]
;       FILING DATE: [illegible] MAY [illegible]
;       APPLICATION NUMBER: [illegible]
;       FILING DATE: [illegible]
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER: [illegible]
;       FILING DATE: [illegible]
;     ATTORNEY/AGENT INFORMATION:
;       NAME: [illegible]
;       REGISTRATION NUMBER: [illegible]
;       REFERENCE/DOCKET NUMBER: [illegible]
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE: [illegible]
;       TELEFAX: [illegible]
;  INFORMATION FOR SEQ ID NO: [illegible]:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH: [illegible]
;       TYPE: amino acid
;       STRANDEDNESS: [illegible]
;       TOPOLOGY: [illegible]
;     MOLECULE TYPE: protein
;     FEATURE:
;       NAME/KEY: [illegible]
;       LOCATION: [illegible]
;       OTHER INFORMATION: [illegible]
US-[illegible]

  Query Match            [illegible]; Score [illegible]; DB [illegible]; Length [illegible];
  Best Local Similarity  [illegible]; Pred. No. [illegible];
  Matches     4;  Conservative     2;  Mismatches     0;  Indels     0;  Gaps     0;

Qy        1 NQEGWK 6
            |:||.:
Db       26 NEEGWR 31


RESULT  15
US-08-373-215-12
; Sequence 12, Application US/08373215
; Patent No. 5973110
;  GENERAL INFORMATION:
;    APPLICANT:  Muller, Daniel
;    APPLICANT:  Brownell, Elise
;    APPLICANT:  Delaria, Katherine
;    APPLICANT:  Wallace, Linda
;    TITLE OF INVENTION:  Cytotoxic T-Lymphocyte Antigen as a
;    TITLE OF INVENTION:  Cysteine Protease Inhibitor
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Banner & Allegretti, Ltd.
;      STREET:  10 S. Wacker Drive Suite 30[illegible]
;      CITY:  Chicago
;      STATE:  IL
;      COUNTRY:  USA
;      ZIP:  60606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/373,215
;      FILING DATE:  18-MAY-1995
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/US9[illegible]
;      FILING DATE:  15-JUL-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/915,92[illegible]
;      FILING DATE:  17-JUL-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McDonnell, John J
;      REGISTRATION NUMBER:  26,949
;      REFERENCE/DOCKET NUMBER:  93613 C
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (312) 715-1000
;      TELEFAX:  (312) 715-1234
;   INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  96 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  not relevant
;      TOPOLOGY:  not relevant
;    MOLECULE TYPE:  protein
;    FEATURE:
;      NAME/KEY:  Protein
;      LOCATION:  1..96
;      OTHER INFORMATION:  /note= "huCAII"
US-08-373-215-12


  Query Match             84.2%;  Score [illegible];  DB 2;  Length 96;
  Best Local Similarity   66.7%;  Prec. No. [illegible];
  Matches     4;  Conservative     2;  Mismatches     0;  Indels     0;  Gaps     0;

Qy        1 NQEGWK 6
            |:|||:
Db       26 NEEGWR 31


Search completed: August 22, 2001, 14:30:12
```

```
Job time: 366 sec
```

```
                          GenCore version 4.5
                 Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on:           August 22, 2001, 14:25:25 ; Search time 64.49 Seconds
                                              (without alignments)
                                              22.442 Million cell updates/sec

Title:            US-08-659-254-5
Perfect score:    103
Sequence:         1 GSTCVLAGMATHFGEMPEQ 19

Scoring table:    BLOSUM62
                  Gapop 10.0 , Gapext 0.5

Searched:         219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters:      219241

  mum DB seq length: 0
  mum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :        PIR_68:*
               1:  pir1:*
               2:  pir2:*
               3:  pir3:*
               4:  pir4:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                                SUMMARIES
                   %
Result           Query
  No.    Score   Match Length DB  ID                        Description
-----------------------------------------------             --------------
     1     103   100.0    284  2  A55455                    steroidogenic acut
     2     103   100.0    362  2  JC5386                    steroidogenic acut
     3      83    80.6    285  2  I38248                    steroidogenic acut
     4      72    69.9    285  2  JC4385                    steroidogenic acut
     5      45    43.7    634  2  C69975                    acyltransferase ho
     6    44.5    43.2    446  2  G75304                    conserved hypothet
     7      44    42.7    440  2  A39613                    oligodendrocyte my
     8      44    42.7   1567  2  T03730                    antigen containing
     9      43    41.7    639  1  C42917                    carbon monoxide  R
    10      43    41.7   1135  2  T30561                    Scythe protein   A
    11      43    41.7   1259  2  T32901                    hypothetical prote
    12      42    40.8    386  2  T16010                    hypothetical prote
    13      42    40.8    439  2  T41966                    hypothetical prote
    14      42    40.8    462  2  B75306                    probable cycloprop
    15      42    40.8   1321  2  T00382                    hypothetical prote
    16      42    40.8   4558  2  C82139                    RTX toxin RtxA Vcl
    17    41.5    40.3    214  2  T51027                    hypothetical prote
    18    41.5    40.3    247  2  T20039                    hypothetical prote
    19    41.5    40.3    477  2  G85968                    probable oxidoredu
    20    41.5    40.3    477  2  C65036                    hypothetical  21.
    21      41    39.8    210  2  H82747                    ubiquinol cytochro
    22      41    39.8    236  2  D69048                    conserved hypothet
    23      41    39.8    252  2  C54938                    cyclase hisF   Rho
    24      41    39.8    269  2  G75158                    hypothetical prote
    25      41    39.8    274  2  F75187                    polyribonucleotide
    26      41    39.8    328  2  A81064                    alanine racemase (
    27      41    39.8    421  2  A54475                    Ztp-57   mouse
    28      41    39.8    429  2  T40112                    hydroxylase family
    29      41    39.8    803  2  T40514                    Chaperonin hsp78p
```

```
    30    40.5    39.3    701  2  T15239                    hypothetical prote
    31      40    38.8    177  2  JC5748                    coronafacic acid s
    32      40    38.8    224  1  B70166                    conserved hypothet
    33      40    38.8    230  2  F19364                    hypothetical prote
    34      40    38.8    274  2  D71132                    probable autoantig
    35      40    38.8    276  2  G82436                    probable thiosulfa
    36      40    38.8    335  2  T03230                    membrane protein,
    37      40    38.8    345  2  S72493                    N-acetyl-gamma-glu
    38      40    38.8    402  1  UHUPI                     plasminogen activa
    39      40    38.8    437  1  F64782                    probable transport
    40      40    38.8    440  2  A47536                    oligodendrocyte my
    41      40    38.8    463  2  G85550                    hypothetical prote
    42      40    38.8    530  2  A70589                    probable etpA prot
    43      40    38.8    580  2  A56382                    purine permease, b
    44      40    38.8    620  2  T30765                    hypothetical prote
    45      40    38.8   1248  2  A47445                    reverse gyrase   S

                               ALIGNMENTS

RESULT    1
A55455
steroidogenic acute regulatory protein precursor, mitochondrial - mouse
C;Species: Mus musculus (house mouse)
C;Date: 10-Feb-1995 #sequence_revision 10-Feb-1995 #text_change 08-Oct-1999
C;Accession: A55455
R;Clark, B.J.; Wells, J.; King, S.R.; Stocco, D.M.
J. Biol. Chem. 269, 28314-28322, 1994
A;Title: The purification, cloning, and expression of a novel luteinizing hormone-ind
ry protein (StAR).
A;Reference number: A55455; MUID:95050[illegible]
A;Accession: A55455
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-284 <CLA>
A;Cross references: GB:L36062; GB:L36061; GB:S[illegible]; NID:g1244242; PIDN:AAB94783.1; P
C;Keywords: mitochondrion; phosphoprotein
F;1-47/Domain: transit peptide (mitochondrion) #status predicted <TRP>

  Query Match             100.0%;  Score 103;  DB 2;  Length 284;
  Best Local Similarity   100.0%;  Pred. No. 7e-09;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GSTCVLAGMATHFGEMPEQ 19
              |||||||||||||||||||
Db        193 GSTCVLAGMATHFGEMPEQ 211

RESULT    2
JC5386
steroidogenic acute regulatory protein - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 04-Jun-1997 #sequence_revision 20-Jun-1997 #text_change 07-May-1999
C;Accession: JC5386
R;Lee, H., Ahn, K.S., Kwon, H.; Soh, J.
Biochem. Biophys. Res. Commun. 230, 528-532, 1997
A;Title: Nucleotide sequence of rat steroidogenic acute regulatory protein complement
A;Reference number: JC5386; MUID:9716[illegible]
A;Accession: JC5386
A;Molecule type: mRNA
A;Residues: 1-362 <LEE>
A;Cross-references: GB:U76419
C;Comment: This protein is a key regulator for the steroidogensis in acute response t

  Query Match             100.0%;  Score 103;  DB 2;  Length 362;
  Best Local Similarity   100.0%;  Pred. No. 8.9e-09;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GSTCVLAGMATHFGEMPEQ 19
              |||||||||||||||||||
Db        271 [illegible]CVLA[illegible]MA[illegible]HF[illegible]MF[illegible]L [illegible]


RESULT      3
I38348
steroidogenic acute regulatory protein - human
C;Species: Homo sapiens (man)
C;Date: 20 May 1998 #sequence_revision [illegible] May 1998 #text_change [illegible]
C;Accession: I38348; I38[illegible]
R;Sugawara, T.; Lin, [illegible] Martin, [illegible]
Biochemistry 34, 12[illegible] 1995
A;Title: Structure of the human steroidogenic acute regulatory protein (StAR) gene: StAR
A;Reference number: I38348; MUID:[illegible]
A;Accession: I38348
A;Status: preliminary; translated from GB/EMBL/[illegible]
A;Molecule type: DNA
A;Residues: 1-285 [illegible]
A;Cross-references: EMBL:[illegible]; NID:[illegible]; PIDN:AA[illegible]
R;Sugawara, T.; Holt, J.A.; Driscoll, D.; Strauss, [illegible]
Proc. Natl. Acad. Sci. U.S.A. [illegible]
A;Title: Human steroidogenic acute regulatory protein [illegible]
A;Reference number: I38806; MUID:[illegible]
A;Accession: I38806
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-285 [illegible]
A;Cross-references: EMBL:[illegible]; NID:g[illegible]; PIDN:AA[illegible]
C;Genetics:
A;Gene: StAR
A;Cross-references: GDB:StAR; GDB:[illegible]; MIM:[illegible]
A;Map position: 8p11 [illegible] 8p[illegible]
A;Introns: 22/1; 60/[illegible] 10[illegible]; 155/[illegible]


  Query Match             [illegible]; Score [illegible]; DB [illegible]; Length 28[illegible];
  Best Local Similarity   [illegible]; Pred. No. [illegible];
  Matches   16; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible];

Qy          1 GSICVLAGMA[illegible]GEM[illegible]
                |:| ||||: |  | : :|
Db        194 GSICVLAGMS[illegible]GNM[illegible]


RESULT      4
JC4315
steroidogenic acute regulatory protein - bovine
C;Species: Bos primigenius taurus (cattle)
C;Date: 29 Nov 1996 #sequence_revision 08 Feb 1997 #text_change [illegible]
C;Accession: JC4315
R;Hartung, S.; Rust, W.; Balvers, M.; [illegible]
Biochem. Biophys. Res. Commun. 215, [illegible]
A;Title: Molecular cloning and in vivo expression [illegible]
A;Reference number: JC4315; MUID:96[illegible]
A;Accession: JC4315
A;Molecule type: mRNA
A;Residues: 1-285 <HAR>
C;Comment: This protein is an acute [illegible]
C;Genetics:
A;Gene: StAR
F;226-264/Region: [illegible]


  Query Match             [illegible]; Score [illegible]; DB [illegible]; Length [illegible];
  Best Local Similarity   [illegible]; Pred. No. [illegible];
  Matches   14; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible];

Qy          1 GSICVLAGMA[illegible]GEM[illegible]
                || ||||| :  ||
Db        194 GSMCVLAGMALHYGEM[illegible]


RESULT      5
S6[illegible]
```

```
acyltransferase [illegible] Bacillus subtilis
C;Species: Bacillus subtilis
C;Date: [illegible] #sequence_revision [illegible] #text_change [illegible]
C;Accession: [illegible]
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Ber[illegible]
[illegible] Carter, N.M.;
[illegible] Errington, J.; Fabret, C.; Ferrari
Nature [illegible]
A;Authors: [illegible]
[illegible] Hilbert, H.; [illegible]
[illegible]
A;Authors: [illegible] Lazarevic, V.; Lee, [illegible] M.;
[illegible] Pohl, T.M.; Portetelle
[illegible]
A;Authors: [illegible] Schneider, K.; Scoffone, [illegible]; Sekiguchi, J.; Sekowska, A.; Se[illegible]
[illegible]
[illegible] Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida
A;Authors: [illegible] H.; Danchin, A.
A;Title: The complete genome sequence of the [illegible]
A;Reference number: A6[illegible]; MUID:98044[illegible]
A;Accession: [illegible]
A;Status: preliminary [illegible]
A;Molecule type: DNA
A;Residues: [illegible]
A;Cross-references: [illegible]; PIDN:CAB1[illegible]
A;Experimental source: [illegible]
C;Genetics:
A;Gene: [illegible]


  Query Match             [illegible]; Score [illegible]; DB [illegible]; Length [illegible];
  Best Local Similarity   [illegible]; Pred. No. [illegible];
  Matches [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible];

Qy          [illegible]VL[illegible]MA[illegible]
Db          [illegible]


RESULT [illegible]
[illegible]
[illegible] (strain R1)
C;Species: [illegible]
C;Date: [illegible] #sequence_revision [illegible] #text_change [illegible] Mar [illegible]
C;Accession: [illegible]
R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, [illegible]
  [illegible] Vamathevan, [illegible] McDonald, [illegible] Utterback, [illegible] Zalewski, [illegible]
  [illegible] Smith, H.[illegible] Venter, [illegible] Fraser, [illegible] M.
Science 286, [illegible]
A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A;Reference number: A7[illegible]; MUID:[illegible]
A;Accession: [illegible]
A;Status: preliminary
A;Molecule type: DNA
A;Residues: [illegible]
A;Cross-references: GB:AE[illegible]; NID:g6[illegible]; PIDN:AAF[illegible]
A;Experimental source: [illegible]
C;Genetics:
A;Gene: DR21[illegible]
A;Map position: [illegible]
C;Superfamily: [illegible]


  Query Match             [illegible]; Score [illegible]; DB [illegible]; Length [illegible];
  Best Local Similarity   [illegible]; Pred. No. [illegible];
  Matches [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible];

Qy          [illegible]
Db          [illegible]


RESULT [illegible]
```

```
A39613
oligodendrocyte-myelin glycoprotein precursor - human
C;Species: Homo sapiens (man)
C;Date: 13-Sep-1991 #sequence_revision [illegible] 1995 #text_change [illegible] Jan 2001
C;Accession: A36688; A34210; A39613; A30187
R;Mikol, D.D.; Alexakos, M.J.; Bayley, C.A.; [illegible]
J. Cell Biol. 111, 2673-2679, 1990
A;Title: Structure and chromosomal localization of the gene for the oligodendrocyte myel
A;Reference number: A36688; MUID:91115958
A;Accession: A36688
A;Molecule type: DNA
A;Residues: 1-440 <MI1>
A;Cross-references: GB:X57436; NID:g23[illegible]; PIDN:CAA40684.1; PID:g23[illegible]
R;Mikol, D.D.; Gulcher, J.R.; Stefansson, K.
J. Cell Biol. 110, 471-479, 1990
A;Title: The oligodendrocyte-myelin glycoprotein belongs to a distinct family of protein
A;Reference number: A34210; MUID:90130636
A;Accession: A34210
A;Molecule type: DNA
A;Residues: 8-20,'X',22-440 <MI2>
A;Cross-references: GB:X51694; NID:g35144; PIDN:CAA35991.1; PID:g35145
R;Viskochil, D.; Cawthon, R.; O'Connell, P.; Xu, G.; Stevens, J.; Culver, M.; Carey, J.;
  Cell. Biol. 11, 906-912, 1991
  tle: The gene encoding the oligodendrocyte-myelin glycoprotein is embedded within th
A;Reference number: A39613; MUID:91117257
A;Accession: A39613
A;Molecule type: mRNA
A;Residues: 1-20,'X',22-440 <VIS>
A;Cross-references: GB:M63623; NID:g189385; PIDN:AAA59970.1; PID:g189386
R;Mikol, D.D.; Stefansson, K.
J. Cell Biol. 106, 1273-1279, 1988
A;Title: A phosphatidylinositol-linked peanut agglutinin-binding glycoprotein in central
A;Reference number: A30187; MUID:88198171
A;Accession: A30187
A;Molecule type: protein
A;Residues: 25-29,'H',31,'H',33-44,'D',46-[illegible],52 <MI3>
C;Genetics:
A;Gene: GDB:OMG; OMGP
A;Cross-references: GDB:127563; OMIM:164345
A;Map position: 17q11.2-17q11.2
A;Introns: #status absent
A;Note: this gene lies within an intron of GDB:NF1 on the opposite strand
C;Superfamily: leucine-rich alpha-2-glycoprotein repeat homology
C;Keywords: blocked carboxyl end; glycoprotein; lipoprotein; phosphatidylinositol linkag
F;1-24/Domain: signal sequence #status predicted <SIG>
F;56-78/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>
F;79-99/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>
F;100-123/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>
F;124-146/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>
F;147-167/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>
  8-191/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>
  2-215/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>
F;216-227/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR8>
F;426-440/Domain: carboxyl-terminal propeptide #status predicted <TIP>
F;425/Modified site: GPI-anchor ethanolamine amidated carboxyl end (Asn) (in mature form


  Query Match             42.7%;  Score 41;  DB 2;  Length 440;
  Best Local Similarity   46.7%;  Pred. No. [illegible];
  Matches    7;  Conservative    4;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          5 VLACMATHFGFMPPQ 19
                : :||  :| |||:|
Db        382 ITSCMPNNFSFMPQQ 396


RESULT   8
T03730
antigen containing epitope to monoclonal antibody MMS-85/12 - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 05-Nov-1999
C;Accession: T03730
R;Fisher, L.W.; Merrelli, A.; Benayahu, D.
submitted to the EMBL Data Library, July 1997
A;Description: Clone containing epitope to mouse osteoblast monoclonal antibody MMS-8
A;Reference number: Z15035
A;Accession: T03730
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1567 <FIS>
A;Cross-references: EMBL:AF013969; NID:g[illegible]; PIDN:AAB69856.1; PID:g2384711


  Query Match             42.7%;  Score 41;  DB 2;  Length 1567;
  Best Local Similarity   50.0%;  Pred. No. [illegible];
  Matches    8;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 GSTCVLACMATHFGEM 16
                || | : | |   ||:
Db        179 GSACTVEGIATFFGEV 194


RESULT   9
C42957
carbon monoxide dehydrogenase (EC 1.2.99.2) beta chain CooS [similarity] - Rhodospiri
C;Species: Rhodospirillum rubrum
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 01-Sep-2000
C;Accession: C42957; [illegible]51321
R;Kerby, R.L.; Hong, S.S.; Ensign, S.A.; Coppoc, L.J.; Ludden, P.W.; Roberts, G.P.
J. Bacteriol. 174, 5284-5294, 1992
A;Title: Genetic and physiological characterization of the Rhodospirillum rubrum carb
A;Reference number: A42957; MUID:92[illegible]
A;Accession: C42957
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: DNA; protein
A;Residues: 1-639 <KER>
A;Cross-references: GB:U65510; GB:[illegible]; GB:[illegible]421; GB:[illegible]41762; GB:U20508; NID:g15154
A;Experimental source: strain UR1
A;Note: sequence extracted from NCBI backbone (NCBIP:[illegible])
C;Genetics:
A;Gene: cooS
C;Superfamily: carbon monoxide dehydrogenase beta chain; hybrid cluster [4Fe-2S-3O] h
C;Keywords: 4Fe-2S-3O; 4Fe-4S; electron transfer; iron-sulfur protein; metalloprotein
F;238-587/Domain: hybrid cluster [4Fe-2S-3O] homology <HCL>
F;50,53,58,72/Binding site: 4Fe-4S cluster (Cys) (covalent) #status predicted
F;265,304,338,451,481,531,566/Binding site: Ni-4Fe-2S cluster (His, Gln, Cys, Cys,
F;451/Modified site: cysteine persulfide (Cys) #status predicted


  Query Match             41.7%;  Score 40;  DB 1;  Length 639;
  Best Local Similarity   47.1%;  Pred. No. [illegible];
  Matches    8;  Conservative    2;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          1 GSTCVLACMATHFGEMP 17
                ||  |  ': || | :|
Db        573 GSWAVI[illegible]HVGSVP 589


RESULT  10
T30561
Scythe protein - African clawed frog
C;Species: Xenopus laevis (African clawed frog)
C;Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 11-May-2000
C;Accession: T30561
R;Thress, K.; Henzel, W.; Shillinglaw, W.; Kornbluth, S.
EMBO J. 17, 6135-6144, 1998
A;Title: Scythe: a novel reaper-binding apoptotic regulator.
A;Reference number: Z20860; MUID:99[illegible]
A;Accession: T30561
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-11[illegible] <THR>
A;Cross-references: EMBL:AF098511; NID:g38[illegible]; PID:g[illegible]8[illegible]139; PIDN:AAC8[illegible]822.1
C;Function:
A;Description: is a apoptotic regulator that is an essential component in the pathway
C;Superfamily: unassigned ubiquitin-related proteins; ubiquitin homology
```

```
  Query Match             1[illegible]
  Best Local Similarity   1[illegible]  Pred. No. [illegible]
  Matches    8; Conservative [illegible]  Mismatches [illegible]

Qy          1 GSICVLAGMAIHEGEMF 1[illegible]
              | |    |   |  |
Db        344 GTIVTAIGNGIHAGEM[illegible] 3[illegible]


RESULT 11
T32901
hypothetical protein T2[illegible] - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision [illegible] #text_change [illegible]
C;Accession: T32901
[illegible]ray, J.; Robinson, [illegible]
[illegible]tted to the EMBL Data Library, January 1998
A;Description: The sequence of C. elegans [illegible]
A;Reference number: [illegible]124
A;Accession: T32901
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1259 <MUR>
A;Cross-references: EMBL:[illegible]; PIDN:AA[illegible]
A;Experimental source: strain Bristol N2; clone [illegible]
C;Genetics:
A;Gene: CESP:[illegible]42[illegible].4
A;Map position: [illegible]
A;Introns: 32/3; 84/1; 16[illegible]; 300/3; [illegible]


  Query Match             1[illegible]
  Best Local Similarity   [illegible]  Pred. No. [illegible]
  Matches    [illegible]; Conservative [illegible]  Mismatches [illegible]

Qy          4 CVLAGMAIHEGMFE[illegible] 1[illegible]
              | |||| |    |
Db        665 C[illegible]GMAYFE[illegible]DEE [illegible]


RESULT 12
T16010
hypothetical protein [illegible] - Caenorhabditis elegans
[illegible]ecies: Caenorhabditis elegans
[illegible]e: 20-Sep-1999 #sequence_revision [illegible] #text_change [illegible]
[illegible]cession: T16010
R;Pauley, A.
submitted to the EMBL Data Library, M[illegible]
A;Description: The sequence of C. elegans [illegible]
A;Reference number: [illegible]14[illegible]
A;Accession: T16010
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-386 <PAU>
A;Cross-references: EMBL:[illegible]; NID:[illegible]; PIDN:AA[illegible]
A;Experimental source: strain Bristol N2
C;Genetics:
A;Gene: CESP:F09E7.4
A;Introns: 19/3; 11[illegible]; 2[illegible]; [illegible]


  Query Match             [illegible]
  Best Local Similarity   [illegible]  Pred. No. [illegible]
  Matches    [illegible]; Conservative [illegible]  Mismatches [illegible]

Qy          1 GSICVLAGMAIHEGEM[illegible] 18
              |    |    |
Db        207 GADAFHAGLAIHEVES[illegible] 2[illegible]


RESULT 13
[illegible]
hypothetical protein [illegible] - human herpesvirus [illegible] strain [illegible])
C;Species: human herpesvirus [illegible]
A;Variety: strain [illegible]
C;Date: [illegible] #sequence_revision [illegible] #text_change [illegible]
C;Accession: [illegible]
R;Ne[illegible]
submitted to the EMBL Data Library, December [illegible]
A;Description: [illegible]
A;Reference number: [illegible]
A;Accession: [illegible]
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: [illegible] <N[illegible]>
A;Cross-references: [illegible]; PIDN:AA[illegible]
A;Experimental source: strain [illegible]
C;Genetics:
A;N[illegible]


  Query Match             [illegible]
  Best Local Similarity   [illegible]  Pred. No. [illegible]
  Matches    [illegible]; Conservative [illegible]  Mismatches [illegible]  Indels [illegible]  Gaps [illegible]

Qy          [illegible]V[illegible]AGMA[illegible]
Db          [illegible]N[illegible]V[illegible]AMA[illegible]


RESULT 14
[illegible]
probable [illegible] fatty [illegible] synthase [illegible]
C;Species: [illegible]
C;Date: [illegible] #sequence_revision [illegible] #text_change [illegible] Mar [illegible]
C;Accession: [illegible]
R;W[illegible]
  M[illegible]; M[illegible]; Vanathe[illegible]; M[illegible]; [illegible]terbank[illegible]; Zalewsk[illegible]
  [illegible]; Venter[illegible]; [illegible]
Science 286, [illegible]
A;Title: [illegible]
A;Reference number: A[illegible]; MUID:[illegible]
A;Accession: [illegible]
A;Status: preliminary
A;Molecule type: DNA
A;Residues: [illegible] <WHI>
A;Cross-references: [illegible]; PIDN:AAF[illegible]
A;Experimental source: [illegible]
C;Genetics:
A;Gene: [illegible]
A;Map [illegible]
C;Superfamily: [illegible] phosphopantetheine [illegible] synthase; MCM homology


  Query Match             [illegible]  Score 42; [illegible]  Length 4[illegible]
  Best Local Similarity   47.[illegible]  Pred. No. [illegible]
  Matches    [illegible]; Conservative [illegible]  Mismatches [illegible]  Indels [illegible]  Gaps [illegible]

Qy          [illegible]V[illegible]AGMA[illegible]M[illegible]
Db          [illegible]AVA[illegible]


RESULT 15
[illegible]
hypothetical protein KIAA0[illegible] - human (fragment)
C;Species: [illegible]
C;Date: [illegible] #sequence_revision [illegible] #text_change 21 Jul 200[illegible]
C;Accession: [illegible]
R;Ishikawa, K.; Nagase, [illegible]; Suyama, M.; Miyajima, N.; Tanaka, A.; Kotani, H.; Nomura,
DNA [illegible]
A;Title: [illegible] sequences of [illegible] unidentified human genes. X. The compl
A;Reference number: [illegible]; MUID:[illegible]
A;Accession: [illegible]
```

```
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-1321 <ISH>
A;Cross-references: EMBL:AB014534; NID:g3327081; PIDN:BAA31609.1; PID:g3327082
A;Experimental source: brain
C;Genetics:
A;Note: KIAA0634


  Query Match             40.8%;  Score 42;  DB 2;  Length 1321;
  Best Local Similarity   50.0%;  Pred. No. [illegible]e+02;
  Matches    9;  Conservative    4;  Mismatches    3;  Indels    2;  Gaps    1;

Qy        1 GSTCVLAGMATHFGEMPE 18
              |::|| || :   ||:||
Db     1115 GTSCVAAGRS--HGEVPE 1130


Search completed: August 22, 2001, 14:25:2
Job time: 81 sec
```

```
                              GenCore version 4.5
                    Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on.          August 22, 2001, 14:30 55 ; Search time 35.8 Seconds
                                             (without alignments)
                                             18.180 Million cell updates/sec

Title:           US-08-659-254-5
Perfect score:   103
Sequence:        1 GSTCVLAGMATHFGEMPEQ 19

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters:        93435

   mum DB seq length: 0
   mum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       SwissProt_39:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                                   SUMMARIES
                   %
Result           Query
   No.   Score   Match Length DB  ID                                    Description
-------------------------------------------------------------------------------------
     1     103   100.0    284  1  STAR_MOUSE                            P51557 mus musculu
     2     103   100.0    284  1  STAR_RAT                              P97826 rattus norv
     3      99    96.1    284  1  STAR_MESAU                            P70114 mesocricet.
     4      89    86.4    285  1  STAR_PIG                              Q28996 sus scrofa
     5      87    84.5    285  1  STAR_HORSE                            O46689 equus cabal
     6      83    80.6    285  1  STAR_HUMAN                            P49675 homo sapien
     7      72    69.9    285  1  STAR_BOVIN                            Q28918 bos taurus
     8      66    64 1    121  1  STAR_SHEEP                            P79245 ovis aries
     9      44    42.7    440  1  OMGP_HUMAN                            P23515 homo sapien
    10      43    41.7    639  1  COOS_RHORU                            P31896 rhodospiril
    11      42    40.8    439  1  UL17_HSV7J                            P52360 human herpe
    12    41.5    40.3    477  1  YGJI_ECOLI                            P42590 escherichia
    13      41    39.8    252  1  HIS6_[illegible]                      [illegible]
    14      41    39.8    358  1  Y4FP_RHISN                            P55454 rhizobium s
    15    40.5    39.3    701  1  OPT2_CAEEL                            Q01840 caenorhabdi
    16      40    38.8    294  1  ARGC_BACST                            Q07906 bacillus st
    17      40    38.8    402  1  PAI1_HUMAN                            P05121 homo sapien
    18      40    38.8    440  1  OMGP_MOUSE                            Q63912 mus musculu
    19      40    38.8    484  1  ALLP_ECOLI                            P75712 escherichia
    20      40    38.8    580  1  UAPC_EMEN                             P48777 emericella
    21      40    38.8   1248  1  TOPG_SULAC                            Q09592 sulfolobus
    22    39.5    38.3    133  1  R24B_SCHPO                            O59665 schizosacch
    23    39.5    38.3    440  1  PPAS_MOUSE                            P35396 mus musculu
    24    39.5    38.3    441  1  PPAS_HUMAN                            Q03181 homo sapien
    25      39    37.9    282  1  YOR3_METJA                            P96800 methanococ
    26      39    37.9    289  1  DHPS_SYNY3                            P73248 synechocyst
    27      39    37.9    321  1  Z177_HUMAN                            Q13360 homo sapien
    28      39    37.9    558  1  Y369_MYCPN                            P75231 mycoplasma
    29      39    37.9    674  1  DCMB_MOOTH                            P27989 moorella th
    30      39    37.9    735  1  CATA_BACST                            P14412 bacillus st
    31      39    37.9    957  1  SECA_MYCSM                            P71533 mycobacteri
    32      39    37.9   1041  1  EGT2_YEAST                            P42835 saccharomyc
    33      39    37.9   2039  1  CR1_HUMAN                             P17927 homo sapien
    34    38.5    37.4    381  1  RFAK_SALTY                            P26473 salmonella
    35    38.5    37.4    396  1  PPAS_XENLA                            P37233 xenopus lae
    36    38.5    37.4    806  1  PARC_BACSU                            Q45066 bacillus su
    37      38    36.9    132  1  YK89_AERPE                            Q9ya49 aeropyrum p
    38      38    36.9    190  1  UCRI_PARDE                            P05417 paracoccus
    39      38    36.9    191  1  UCRI_RHOCA                            P08500 rhodobacter
    40      38    36.9    200  1  Z266_HUMAN                            Q14584 homo sapien
    41      38    36.9    247  1  YCP4_YEAST                            P25349 saccharomyc
    42      38    36.9    259  1  CFAD_MOUSE                            P03953 mus musculu
    43      38    36.9    263  1  CFAD_RAT                              P32038 rattus norv
    44      38    36.9    309  1  YQGH_BACSU                            P46339 bacillus su
    45      38    36.9    404  1  EAD_EBV                               P03191 epstein bar


                                   ALIGNMENTS


RESULT    1
STAR_MOUSE
ID   STAR_MOUSE      STANDARD;      PRT;   284 AA.
AC   P51557;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN PRECURSOR (LUTEINIZING HORMONE-
DE   INDUCED PROTEIN).
GN   STAR.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX   MEDLINE=95050616; PubMed=7961773;
RA   Clark B.J., Wells J., King S.R., Stocco D.M.;
RT   "The purification, cloning, and expression of a novel luteinizing
RT   hormone-induced mitochondrial protein in MA 10 mouse Leydig tumor
RT   cells. Characterization of the steroidogenic acute regulatory protein
RT   (StAR).";
RL   J. Biol. Chem. 269:28314-28322(1994).
CC   -!- FUNCTION: PLAYS A KEY ROLE IN STEROID HORMONE SYNTHESIS BY
CC        ENHANCING THE METABOLISM OF CHOLESTEROL INTO PREGNENOLONE.
CC        TRANSPORTER THAT BINDS TO AND TRANSPORT CHOLESTEROL THROUGH THE
CC        INTERMEMBRANE SPACE OF THE MITOCHONDRION (PROBABLE).
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL.
CC   -!- SIMILARITY: BELONGS TO THE STAR FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; L36062; AAB94783.1; -.
DR   MGD; MGI:102760; Star.
DR   InterPro; IPR000799; -.
DR   InterPro; IPR002913; -.
DR   Pfam; PF01852; START; 1.
DR   PRINTS; PR00978; STARPROTEIN.
KW   Lipid binding; Lipid transport; Transport; Steroidogenesis;
KW   Mitochondrion; Transit peptide.
FT   TRANSIT       1      ?       MITOCHONDRION (POTENTIAL).
FT   CHAIN         ?    284       STEROIDOGENIC ACUTE REGULATORY PROTEIN
FT   DOMAIN      170    174       POLY-ALA.
SQ   SEQUENCE   284 AA;  31626 MW;  9B3 0E5A144 9B72 C064;


  Query Match             100.0%;  Score 103;  DB 1;  Length 284;
  Best Local Similarity   100.0%;  Pred. No. 2.2e-09;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy        1 GSIQVLAGMAIHEGEMPEQ 19
            [illegible]
Db      193 GSIQVLAGMAIHEGEMPEQ 211


RESULT 2
STAR RAT
ID   STAR RAT      STANDARD;      PRT;   284 AA.
AC   P97826; P97[illegible]
DT   01 NOV 1997 (Rel. 35, Created)
DT   01 NOV 1997 (Rel. 35, Last sequence update)
DT   30 MAY 2000 (Rel. 39, Last annotation update)
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN [illegible]
GN   STAR.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; [illegible]
OC   Mammalia; Eutheria; Rodentia; [illegible] Muridae; Murinae; Rattus.
OX   NCBI TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=WISTAR; [illegible]
RX   MEDLINE=97466089; PubMed=[illegible]
RA   Mizutani T., [illegible] Miyamoto K.;
RT   "Molecular cloning, characterization and cellular [illegible]
RT   steroidogenic acute regulatory protein (StAR) [illegible]"
RL   Life Sci. 61:1497-1506(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY; TISSUE=[illegible]
RX   MEDLINE=97167737; PubMed=[illegible]
RA   Lee H.K., Ahn R., Kwon H.B., [illegible]
RT   "Nucleotide sequence of rat steroidogenic [illegible]
RT   complementary DNA.";
RL   Biochem. Biophys. Res. Commun. [illegible]
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY; TISSUE=Adrenal [illegible]
RX   MEDLINE=98184877; PubMed=[illegible]
RA   Ariyoshi N., Kim Y.C., Artemenko I., Bhattacharyya K.K.,
RA   Jefcoate C.R.;
RT   "Characterization of the rat Star gene [illegible]
RT   3.5 kilobase pair mRNA. ACTH [illegible]
RT   precedes elevation of Star mRNA and protein."
RL   J. Biol. Chem. [illegible](1998).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY; TISSUE=Adrenal gland;
RX   MEDLINE=99048[illegible]; PubMed=[illegible]
RA   Furukawa A., Miyatake A., Ohnishi T., Ichikawa Y.;
RT   "Steroidogenic acute regulatory protein (StAR) [illegible]
RT   constitutively expressed in the adult rat central [illegible]
RT   colocalization of StAR [illegible]
RT   hydroxysteroid dehydrogenase in the rat brain."
RL   J. Neurochem. 71:2231-2238(1998).
CC   -!- FUNCTION: PLAYS A KEY ROLE IN STEROID HORMONE SYNTHESIS BY
CC       ENHANCING THE METABOLISM OF CHOLESTEROL INTO PREGNENOLONE.
CC       TRANSPORTER THAT BINDS TO AND TRANSPORT CHOLESTEROL THROUGH THE
CC       INTERMEMBRANE SPACE OF THE MITOCHONDRION (PROBABLE).
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL.
CC   -!- SIMILARITY: BELONGS TO THE STAR FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AB006349; BAA19[illegible]
DR   EMBL; U76419; AAB[illegible]
DR   EMBL; AF044083; AAC[illegible]
DR   EMBL; AB0[illegible]; BAA[illegible]
```

```
DR   InterPro; IPR[illegible]
DR   InterPro; IPR[illegible]
DR   Pfam; PF[illegible]; START; 1.
DR   PRINTS; PR00978; STARPROTEIN.
KW   Lipid-binding; Lipid transport; Transport; Steroidogenesis;
KW   Mitochondrion; Transit peptide.
FT   TRANSIT       1     [illegible]       MITOCHONDRION (POTENTIAL).
FT   CHAIN         [illegible]  284       STEROIDOGENIC ACUTE REGULATORY PROTEIN.
FT   CONFLICT      [illegible]            MLLATFK -> MLWKSTAVESLVYV[illegible]
FT                                        [illegible]YPGLIQTEQ[illegible]MEITENLEISQGLVVVNHYLAPQLFAKK
FT                                        YYEPSEV AY[illegible] (IN REF. [illegible]).
SQ   SEQUENCE   284 AA;  [illegible] MW;  [illegible] CRC64;

  Query Match          [illegible]%;  Score [illegible];  DB [illegible];  Length 284;
  Best Local Similarity   [illegible]%;  Pred. No. [illegible];
  Matches   [illegible];  Conservative   [illegible];  Mismatches   [illegible];  Indels   [illegible];  Gaps   [illegible];

Qy        [illegible] VLAGMA[illegible]EMPE[illegible]
            [illegible]
Db      [illegible] VLAGMAIHEGEMPE[illegible]


RESULT 3
STAR MESAU
ID   STAR MESAU     STANDARD;      PRT;   284 AA.
AC   [illegible]
DT   [illegible] NOV [illegible] (Rel. [illegible], Created)
DT   [illegible] NOV [illegible] (Rel. [illegible], Last sequence update)
DT   [illegible] MAY [illegible] (Rel. [illegible], Last annotation update)
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN PRECURSOR.
GN   STAR.
OS   Mesocricetus auratus (Golden hamster).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC   Mesocricetus.
OX   NCBI TaxID=[illegible];
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=[illegible]; PubMed=[illegible]
RA   Fleury A., [illegible] M., [illegible] A., Lehoux J.;
RA   [illegible]
RT   "Adrenocorticotropin regulates the level of the steroidogenic acute
RT   regulatory (STAR) protein mRNA in hamster adrenals.";
RL   Endocr. Res. 22:[illegible](1996).
CC   -!- FUNCTION: PLAYS A KEY ROLE IN STEROID HORMONE SYNTHESIS BY
CC       ENHANCING THE METABOLISM OF CHOLESTEROL INTO PREGNENOLONE.
CC       TRANSPORTER THAT BINDS TO AND TRANSPORT CHOLESTEROL THROUGH THE
CC       INTERMEMBRANE SPACE OF THE MITOCHONDRION (PROBABLE).
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL.
CC   -!- SIMILARITY: BELONGS TO THE STAR FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; [illegible]; AAB[illegible]
DR   InterPro; IPR00[illegible]
DR   InterPro; IPR[illegible]
DR   Pfam; PF[illegible]; START; 1.
DR   PRINTS; PR00978; STARPROTEIN.
KW   Lipid-binding; Lipid transport; Transport; Steroidogenesis;
KW   Mitochondrion; Transit peptide.
FT   TRANSIT       1     [illegible]       MITOCHONDRION (POTENTIAL).
FT   CHAIN         [illegible]  284       STEROIDOGENIC ACUTE REGULATORY PROTEIN.
SQ   SEQUENCE   284 AA;  [illegible] MW;  [illegible] CRC64;

  Query Match          [illegible];  Length 284;
  Best Local Similarity   94.7%;  Pred. No. [illegible];
  Matches   18;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 GSTCVLAGMATHEGEMPEQ 19
            ||||||||:||||||||||
Db      193 GSTCVLAGIATHEGEMPEQ 211


RESULT    4
STAR_PIG
ID   STAR_PIG        STANDARD;      PRT;   285 AA.
AC   Q28996; Q95259;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN PRECURSOR.
GN   STAR.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Ovary;
RX   MEDLINE=97200784; PubMed=9048613;
RA   Pilon N., Daneau I., Brisson C., Ethier J.F., Lussier J.G.,
RA   Silversides D.W.;
RT   "Porcine and bovine steroidogenic acute regulatory protein (StAR)
RT   gene expression during gestation.";
RL   Endocrinology 138:1085-1091(1997).
RN   [2]
RP   SEQUENCE OF 132-210 FROM N.A.
RC   TISSUE=Corpus luteum;
RX   MEDLINE=97131976; PubMed=8977433;
RA   Balasubramanian K., Lavoie H.A., Garmey J.C., Stocco D.M.,
RA   Veldhuis J.D.;
RT   "Regulation of porcine granulosa cell steroidogenic acute regulatory
RT   protein (StAR) by insulin-like growth factor I: synergism with
RT   follicle-stimulating hormone or protein kinase A agonist.";
RL   Endocrinology 138:433-439(1997).
CC   -!- FUNCTION: PLAYS A KEY ROLE IN STEROID HORMONE SYNTHESIS BY
CC       ENHANCING THE METABOLISM OF CHOLESTEROL INTO PREGNENOLONE.
CC       TRANSPORTER THAT BINDS TO AND TRANSPORT CHOLESTEROL THROUGH THE
CC       INTERMEMBRANE SPACE OF THE MITOCHONDRION (PROBABLE).
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIA.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN ADULT AND FETAL STEROIDOGENIC
CC       TISSUES, INCLUDING ADULT TESTES AND OVARIES AND ADULT ADRENAL
CC       GLANDS AS WELL AS STEROIDOGENIC TISSUES OF PREGNANCY, INCLUDING
CC       DEVELOPING FETAL TESTES, CORPUS LUTEUM, AND PREGNANCY, BUT NOT IN
CC       THE FETAL OVARY.
CC   -!- SIMILARITY: BELONGS TO THE STAR FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; U53020; AAB04953.1; -.
DR   EMBL; U72195; AAB50555.1; -.
DR   InterPro; IPR000799; -.
DR   InterPro; IPR002913; -.
DR   Pfam; PF01852; START; 1.
DR   PRINTS; PR00978; STARPROTEIN.
KW   Lipid-binding; Lipid transport; Transport; Steroidogenesis;
KW   Mitochondrion; Transit peptide.
FT   TRANSIT       1      ?       MITOCHONDRION (POTENTIAL).
FT   CHAIN         ?    285       STEROIDOGENIC ACUTE REGULATORY PROTEIN.
FT   CONFLICT    188    188       G -> R (IN REF. 2).
SQ   SEQUENCE   285 AA;  31833 MW;  67DCA[illegible]21519B8 CRC64;

  Query Match             86.4%;  Score 49;  DB 1;  Length 285;
  Best Local Similarity   89.5%;  Pred. No. 3.9e-07;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy        1 GSTCVLAGMATHEGEMPEQ 19
            || |||||| || ||||||
Db      194 GSVCVLAGMATDEGEMPEQ 212


RESULT    5
STAR_HORSE
ID   STAR_HORSE      STANDARD;      PRT;   285 AA.
AC   O46689;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN PRECURSOR.
GN   STAR.
OS   Equus caballus (Horse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX   NCBI_TaxID=9796;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Follicle;
RX   MEDLINE=99124482; PubMed=9927292;
RA   Kerban A., Boerboom D., Sirois J.;
RT   "Human chorionic gonadotropin induces an inverse regulation of
RT   steroidogenic acute regulatory protein messenger ribonucleic acid in
RT   theca interna and granulosa cells of equine preovulatory follicles.";
RL   Endocrinology 140:667-674(1999).
CC   -!- FUNCTION: PLAYS A KEY ROLE IN STEROID HORMONE SYNTHESIS BY
CC       ENHANCING THE METABOLISM OF CHOLESTEROL INTO PREGNENOLONE.
CC       TRANSPORTER THAT BINDS TO AND TRANSPORT CHOLESTEROL THROUGH THE
CC       INTERMEMBRANE SPACE OF THE MITOCHONDRION (PROBABLE).
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL.
CC   -!- SIMILARITY: BELONGS TO THE STAR FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AF031697; AAC04704.1; -.
DR   EMBL; AF031696; AAC04703.1; -.
DR   InterPro; IPR000799; -.
DR   InterPro; IPR002913; -.
DR   Pfam; PF01852; START; 1.
DR   PRINTS; PR00978; STARPROTEIN.
KW   Lipid-binding; Lipid transport; Transport; Steroidogenesis;
KW   Mitochondrion; Transit peptide.
FT   TRANSIT       1      ?       MITOCHONDRION (POTENTIAL).
FT   CHAIN         ?    285       STEROIDOGENIC ACUTE REGULATORY PROTEIN.
SQ   SEQUENCE   285 AA;  [illegible] MW;  [illegible]727D CRC64;

  Query Match             84.5%;  Score [illegible];  DB 1;  Length 285;
  Best Local Similarity   89.5%;  Pred. No. [illegible];
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy        1 GSTCVLAGMATHEGEMPEQ 19
            |||||||||||  |||||
Db      194 GSTCVLAGMATQEEEMPEQ 212


RESULT    6
STAR_HUMAN
ID   STAR_HUMAN      STANDARD;      PRT;   285 AA.
AC   P49675; Q16[illegible];
```

```
DT   01 FEB 1996 (Rel. 3[illegible] Created)
DT   01 FEB 1996 (Rel. 3[illegible] Last sequence update)
DT   30 MAY 2000 (Rel. 39, Last annotation update)
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN [illegible]
GN   STAR.
OS   Homo sapiens (Human)
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; [illegible]
OC   Mammalia; Eutheria; Primates; Catarrhini; [illegible]
OX   NCBI_TaxID=[illegible]
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Adrenal cortex;
RX   MEDLINE=95287[illegible]; PubMed=[illegible]
RA   Sugawara T., Holt J.A., Driscoll D., [illegible]
RA   Miller W.L., Pitters [illegible]
RA   Stocco D.M.;
RT   "Human steroidogenic [illegible]
RT   COS-1 cells, tissue-specific expression, and [illegible]
RT   structural gene to 8p11.2 and a pseudogene [illegible]"
RL   Proc. Natl. Acad. Sci. U.S.A. [illegible]
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Placenta;
RX   MEDLINE=96038[illegible]; PubMed=[illegible]
RA   Sugawara T., Lin D., Holt J.A., Martin K. [illegible] W.
RA   Strauss J.F. III;
RT   "Structure of the human steroidogenic acute regulatory [illegible] (STAR
RT   gene: STAR stimulates mitochondrial [illegible]
RT   activity.";
RL   Biochemistry 34:12506-12512(1995).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96008[illegible]; PubMed=[illegible]
RA   Gradi A., Tang W. [illegible] McBride H.M. [illegible]
RA   Pelletier J.;
RT   "The human steroidogenic acute regulatory (StAR) [illegible]
RT   the urogenital system and [illegible]"
RL   Biochim. Biophys. Acta 1258:228-233(1995).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RA   Yu W., Sarqusen [illegible] R.A.
RL   Submitted (NOV-1997) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   VARIANT CLAH LEU-182.
RX   MEDLINE=97096837; PubMed=[illegible]
RA   Bose H.S., Sugawara T., Strauss J.F. III, Miller W.
RT   "The pathophysiology and genetics of [illegible]
RT   hyperplasia.";
RL   New Engl. J. Med. 335:1870-1878(1996).
RN   [6]
RP   VARIANT ALA-[illegible]
RX   MEDLINE=98112[illegible]; PubMed=[illegible]
RA   Katsumata N., Tanae A., Shinagawa [illegible] Nakagomi M. [illegible]
RA   Shimizu M., Yasunaga [illegible] Tanaka T. [illegible]
RT   "A novel frameshift mutation 840delA and [illegible]
RT   the steroidogenic acute regulatory protein gene [illegible]
RT   with congenital lipoid adrenal hyperplasia.";
RL   Hum. Mutat. Suppl. 1:S[illegible](1998).
CC   -!- FUNCTION: PLAYS A KEY ROLE IN STEROID HORMONE SYNTHESIS BY
CC       ENHANCING THE METABOLISM OF CHOLESTEROL INTO PREGNENOLONE.
CC       TRANSPORTER THAT BINDS TO AND TRANSFERS CHOLESTEROL [illegible]
CC       INTERMEMBRANE SPACE OF THE MITOCHONDRION [illegible]
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN GONADS, ADRENAL [illegible]
CC       KIDNEY.
CC   -!- DISEASE: DEFECTS IN STAR CAUSE CONGENITAL LIPOID ADRENAL
CC       HYPERPLASIA (CLAH), THE MOST SEVERE FORM OF ADRENAL [illegible]
CC   -!- SIMILARITY: BELONGS TO THE STAR FAMILY.
CC   --
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics [illegible]
CC   the European Bioinformatics Institute. There are no [illegible]
```

```
CC   [illegible] was
CC   [illegible] statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --
DR   EMBL; [illegible]; AAC[illegible]
DR   EMBL; [illegible]; AAC[illegible]
DR   EMBL; [illegible]; AAC[illegible]; -. [illegible]
DR   EMBL; [illegible]; AAC[illegible]; -. [illegible]
DR   EMBL; [illegible]; AAC[illegible]; -. [illegible]
DR   EMBL; [illegible]; AAC[illegible]; -. [illegible]
DR   EMBL; [illegible]; AAC[illegible]; -. [illegible]
DR   EMBL; [illegible]; AAC[illegible]; -. [illegible]
DR   EMBL; [illegible]; AAB[illegible]
DR   EMBL; [illegible]; AAB[illegible]
DR   MIM; [illegible]
DR   MIM; [illegible]
DR   InterPro; IPR[illegible]
DR   InterPro; IPR[illegible]
DR   Pfam; PF[illegible]; START; 1.
DR   PROSITE; [illegible]
KW   Lipid-binding; Lipid transport; Transport; Steroidogenesis;
KW   Mitochondrion; Transit peptide; Disease mutation; Polymorphism.
FT   TRANSIT      [illegible]      MITOCHONDRION (POTENTIAL).
FT   CHAIN        [illegible]      STEROIDOGENIC ACUTE REGULATORY PROTEIN.
FT   VARIANT      [illegible]      [illegible] (IN CLAH).
FT                                 /FTId=VAR_[illegible]
FT   VARIANT      [illegible]      [illegible] A.
FT                                 /FTId=VAR_[illegible]
SQ   SEQUENCE   [illegible] AA;  [illegible] MW;  [illegible] CRC64;

  Query Match  [illegible]  Score [illegible]  Length 285;
  Best Local Similarity [illegible]  [illegible]
  Matches [illegible]  Conservative [illegible]  Mismatches [illegible]  Indels [illegible]  Gaps [illegible]

Qy   [illegible] VAGMAHEEMEQ [illegible]
               [illegible]
Db   [illegible] VAGMAHEEMEQ [illegible]


RESULT [illegible]
STAR_BOVIN
ID   STAR_BOVIN     STANDARD;      PRT;   285 AA.
AC   [illegible]
DT   [illegible] NOV [illegible] (Rel. [illegible] Created)
DT   [illegible] FEB [illegible] (Rel. [illegible] Last sequence update)
DT   [illegible] MAY [illegible] (Rel. [illegible] Last annotation update)
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN PRECURSOR
GN   STAR.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=[illegible]
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Corpus luteum;
RX   MEDLINE=[illegible]; PubMed=[illegible]
RA   Hartung S., Rust W., Balvers M., Ivell R.;
RT   "M[illegible] expression of the bovine steroidogenic
RT   acute regulatory protein.";
RL   [illegible]
RN   [2]
RP   REVISION [illegible] AND [illegible]
RA   Hartung S., Rust W., Balvers M., Ivell R.;
RL   Submitted (MAY-[illegible]) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: PLAYS A KEY ROLE IN STEROID HORMONE SYNTHESIS BY
CC       ENHANCING THE METABOLISM OF CHOLESTEROL INTO PREGNENOLONE.
CC       TRANSPORTER THAT BINDS [illegible] AND TRANSFERS CHOLESTEROL THROUGH THE
CC       INTERMEMBRANE SPACE OF THE MITOCHONDRION (PROBABLE).
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL.
CC   -!- TISSUE SPECIFICITY: CORPUS LUTEUM AND ADRENAL GLAND.
CC   -!- SIMILARITY: BELONGS TO THE STAR FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; Y17259; CAA76717.1; -.
DR   EMBL; Y17260; CAA76718.1; -.
DR   InterPro; IPR000799; -.
DR   InterPro; IPR002913; -.
DR   Pfam; PF01852; START; 1.
DR   PRINTS; PR00978; STARPROTEIN.
KW   Lipid-binding; Lipid transport; Transport; Steroidogenesis;
KW   Mitochondrion; Transit peptide.
FT   TRANSIT       1      ?       MITOCHONDRION (POTENTIAL).
FT   CHAIN         ?    285       STEROIDOGENIC ACUTE REGULATORY PROTEIN.
FT   CONFLICT     75     75       H -> Y (IN GENOMIC SEQUENCE).
FT   CONFLICT     99     99       E -> K (IN GENOMIC SEQUENCE).
SQ   SEQUENCE   285 AA;  31845 MW;  EAFD5C14E895E235 CRC64;

  Query Match             69.9%;  Score 72;  DB 1;  Length 285;
  Best Local Similarity   73.7%;  Pred. No. [illegible];
  Matches   14;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 GSTCVLAGMATEFGEMPEQ 19
              || |||||||| : |||:|
Db        194 GSMCVLAGMATLYEEMPQQ 212


RESULT   8
STAR_SHEEP
ID   STAR_SHEEP     STANDARD;      PRT;   121 AA.
AC   P79245;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN (FRAGMENT).
GN   STAR.
OS   Ovis aries (Sheep).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Caprinae; Ovis.
OX   NCBI_TaxID=9940;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Corpus luteum;
RX   MEDLINE=96079887; PubMed=7588291;
RA   Juengel J.L., Meberg B.M., Turzillo A.M., Nett T.M., Niswender G.D.;
RT   "Hormonal regulation of messenger ribonucleic acid encoding
RT   steroidogenic acute regulatory protein in ovine corpora lutea.";
RL   Endocrinology 136:5423-5429(1995).
CC   -!- FUNCTION: PLAYS A KEY ROLE IN STEROID HORMONE SYNTHESIS BY
CC       ENHANCING THE METABOLISM OF CHOLESTEROL INTO PREGNENOLONE.
CC       TRANSPORTER THAT BINDS TO AND TRANSPORT CHOLESTEROL THROUGH THE
CC       INTERMEMBRANE SPACE OF THE MITOCHONDRION (PROBABLE).
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL.
CC   -!- SIMILARITY: BELONGS TO THE STAR FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; S80098; AAB47088.1; -.
DR   InterPro; IPR000799; -.
```

```
DR   PRINTS; PR00978; STARPROTEIN.
KW   Lipid-binding; Lipid transport; Transport; Steroidogenesis;
KW   Mitochondrion.
FT   NON_TER       1      1
SQ   SEQUENCE   121 AA;  13667 MW;  E85A[illegible]25EE CRC64;

  Query Match             64.1%;  Score 66;  DB 1;  Length 121;
  Best Local Similarity   68.4%;  Pred. No. 0.0087;
  Matches   13;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 GSTCVLAGMATEFGEMPEQ 19
              || ||||| |  : |||:|
Db         98 GSMCVLAGTAILYEEMPQQ 116


RESULT   9
OMGP_HUMAN
ID   OMGP_HUMAN     STANDARD;      PRT;   440 AA.
AC   P23515;
DT   01-NOV-1991 (Rel. 20, Created)
DT   01-MAY-1992 (Rel. 22, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   OLIGODENDROCYTE-MYELIN GLYCOPROTEIN PRECURSOR
GN   OMG OR OMGP.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91117257; PubMed=1899288;
RA   Viskochil D., Cawthon R.M., O'Connell P., Xu G., Stevens J.,
RA   Culver M., Carey J., White R.;
RT   "The gene encoding the oligodendrocyte-myelin glycoprotein is
RT   embedded within the neurofibromatosis type 1 gene.";
RL   Mol. Cell. Biol. 11:906-912(1991)
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90133635; PubMed=1688857;
RA   Mikol D.D., Gulcher J., Stefansson K.;
RT   "The oligodendrocyte myelin glycoprotein belongs to a distinct family
RT   of proteins and contains the HNK-1 carbohydrate.";
RL   J. Cell Biol. 110:471-479(1990).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91115958; PubMed=2277079;
RA   Mikol D.D., Alexakos M.J., Bayley C.A., Lemons R.S., Le Beau M.M.,
RA   Stefansson K.;
RT   "Structure and chromosomal localization of the gene for the
RT   oligodendrocyte-myelin glycoprotein.";
RL   J. Cell Biol. 111:2673-2679(1990)
RN   [4]
RP   SEQUENCE OF 25-52.
RX   MEDLINE=88198377; PubMed=3283151;
RA   Mikol D.D., Stefansson K.;
RT   "A phosphatidylinositol linked peanut agglutinin binding glycoprotein
RT   in central nervous system myelin and on oligodendrocytes.";
RL   J. Cell Biol. 106:1273-1279(1988).
CC   -!- FUNCTION: CELL ADHESION MOLECULE CONTRIBUTING TO THE INTERACTIVE
CC       PROCESS REQUIRED FOR MYELINATION IN THE CENTRAL NERVOUS SYSTEM.
CC   -!- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR.
CC   -!- TISSUE SPECIFICITY: OLIGODENDROCYTES AND MYELIN OF THE CENTRAL
CC       NERVOUS SYSTEM.
CC   -!- PTM: O-GLYCOSYLATED IN ITS SER/THR-RICH REPEAT DOMAIN (POTENTIAL).
CC   -!- SIMILARITY: CONTAINS 8 LEUCINE-RICH REPEATS (LRR).
CC   -!- CAUTION: DO NOT CONFUSE OLIGODENDROCYTE-MYELIN GLYCOPROTEIN (OMG)
CC       WITH MYELIN OLIGODENDROCYTE GLYCOPROTEIN (MOG).
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed. [illegible]
CC   entities requires a license agreement (See http://[illegible]/announce/
CC   or send an email to [illegible])
CC   [illegible]
DR   EMBL; M63623; AAA59[illegible]; [illegible]
DR   EMBL; X57436; CAA40[illegible]; -.
DR   EMBL; X51694; CAA35[illegible]; [illegible]
DR   PIR; A30187; A30187.
DR   PIR; A34210; A34210.
DR   PIR; A36688; A36688.
DR   PIR; A39613; A39613.
DR   MIM; 164345; -.
DR   InterPro; IPR00[illegible]
DR   InterPro; IPR00[illegible]
DR   Pfam; PF0056[illegible]; LRR; [illegible]
DR   Pfam; PF01462; LRRNT; 1.
DR   PRINTS; PR000[illegible]; LEU[illegible]
KW   Glycoprotein; Myelin; Cell adhesion; Repeat; [illegible]
KW   Signal; GPI-anchor.
FT   SIGNAL        1     [illegible]
FT   CHAIN        [illegible]   4[illegible]       [illegible]
FT   REPEAT       [illegible]   [illegible]        LRR 1.
FT   REPEAT       [illegible]   [illegible]        LRR 2.
FT   REPEAT       [illegible]   [illegible]        LRR 3.
FT   REPEAT       [illegible]   [illegible]        LRR 4.
FT   REPEAT       145    [illegible]       LRR 5.
FT   REPEAT       [illegible]   [illegible]        LRR 6.
FT   REPEAT       190    2[illegible]       LRR 7.
FT   REPEAT       [illegible]   [illegible]        LRR 8 (INCOMPLETE).
FT   REPEAT       [illegible]   [illegible]        SER/THR-RICH.
FT   REPEAT       [illegible]   [illegible]        SER/THR-RICH.
FT   REPEAT       [illegible]   [illegible]        SER/THR-RICH.
FT   REPEAT       [illegible]   [illegible]        SER/THR-RICH.
FT   REPEAT       [illegible]   4[illegible]       SER/THR-RICH.
FT   CARBOHYD     45     [illegible]       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD     [illegible]   [illegible]        N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD     103    10[illegible]     N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD     152    15[illegible]     N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD     1[illegible]   1[illegible]      N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD     18[illegible]  18[illegible]     N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD     [illegible]   1[illegible]       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD     234    2[illegible]      N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD     3[illegible]   3[illegible]      N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD     38[illegible]  38[illegible]     N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD     4[illegible]   4[illegible]      N-LINKED (GLCNAC...) (POTENTIAL).
SQ   SEQUENCE   440 AA;  4[illegible] MW;  [illegible] CRC64;

  Query Match             [illegible]%;  Score 11;  DB [illegible];  Length [illegible];
  Best Local Similarity   [illegible]%;  Pred. No. [illegible];
  Matches    [illegible];  Conservative    4;  Mismatches    [illegible];  Indels  [illegible];  Gaps  [illegible];

Qy          5 VLAGMATHE[illegible]EMREQ [illegible]
              [illegible]
Db        382 [illegible]MPNN[illegible]EMREQ [illegible]


RESULT 10
COOS_RHORU
ID   COOS_RHORU     STANDARD;      PRT;   [illegible] AA.
AC   P31896;
DT   01-JUL-1993 (Rel. 26, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   CARBON MONOXIDE DEHYDROGENASE ([illegible])
GN   COOS.
OS   Rhodospirillum rubrum.
OC   Bacteria; Proteobacteria; alpha subdivision; [illegible]
OC   Rhodospirillum.
OX   NCBI_TaxID=1085;
RN   [1]
RP   SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
RC   STRAIN=UR2;
```

```
RX   MEDLINE=[illegible]; PubMed=[illegible];
RA   Kerby R.L., Hong S.S., Ensign S.A., Coppoc L.J., Ludden P.W.,
RA   [illegible]
RT   "Genetic and physiological characterization of the Rhodospirillum
RT   rubrum carbon monoxide dehydrogenase system.";
RL   J. Bacteriol. 174:5284-5294(1992).
RN   [illegible]
RP   SEQUENCE FROM N.A.
RC   STRAIN=[illegible]
RX   MEDLINE=[illegible]; PubMed=[illegible];
RA   [illegible] Kerby R.L., [illegible] Ludden P.W.;
RT   "Characterization of the CO-induced, CO-tolerant hydrogenase from
RT   Rhodospirillum rubrum and the gene encoding the large subunit of the
RT   enzyme.";
RL   J. Bacteriol. [illegible]
CC   -!- FUNCTION: [illegible] CARBON MONOXIDE [illegible] VIA [illegible]
CC       [illegible] BY AN [illegible] (PROBABLY [illegible]);
CC   -!- CATALYTIC ACTIVITY: [illegible] H(2)O + A [illegible] (2) + REDUCED
CC       ACCEPTOR.
CC   -!- COFACTOR: IRON, NICKEL AND ZINC.
CC   -!- INDUCTION: [illegible] CARBON MONOXIDE [illegible]
CC   -!- MISCELLANEOUS: METHYL VIOLOGEN CAN ACT AS ACCEPTOR.
CC   -!- SIMILARITY: [illegible] HERMIATELLUM CARBON MONOXIDE
CC       DEHYDROGENASE [illegible] AND SOME [illegible] CARBON
CC       MONOXIDE DEHYDROGENASE ALPHA SUBUNIT.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to [illegible])
CC   --------------------------------------------------------------------------
DR   EMBL; [illegible]; AAA[illegible]
DR   PIR; [illegible]
KW   [illegible]
FT   METAL        [illegible]   [illegible]        IRON-SULFUR (BY SIMILARITY).
FT   METAL        [illegible]   [illegible]        IRON-SULFUR (BY SIMILARITY).
FT   METAL        [illegible]   [illegible]        IRON-SULFUR (BY SIMILARITY).
SQ   SEQUENCE   [illegible] AA;  [illegible] MW;  [illegible] CRC64;

  Query Match             [illegible]%;  Score 4[illegible];  DB 1;  Length [illegible];
  Best Local Similarity   [illegible]%;  Pred. No. [illegible];
  Matches    [illegible];  Conservative    [illegible];  Mismatches    [illegible];  Indels  [illegible];  Gaps  [illegible];

Qy          [illegible]VLAGMATHE[illegible]EM[illegible]
              [illegible]
Db        [illegible]SWAV[illegible]


RESULT [illegible]
UL[illegible]_HSV[illegible]
ID   [illegible]_HSV[illegible]     STANDARD;      PRT;   4[illegible] AA.
AC   P[illegible];
DT   [illegible] (Rel. [illegible], Created)
DT   [illegible] (Rel. [illegible], Last sequence update)
DT   [illegible] (Rel. [illegible], Last annotation update)
DE   PROTEIN [illegible]
GN   [illegible]
OS   Human herpesvirus [illegible] (HHV[illegible]).
OC   Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC   Alphaherpesvirinae; [illegible]
OX   NCBI_TaxID=[illegible];
RN   [1]
RP   SEQUENCE FROM N.A.
RA   [illegible]
RL   Submitted [illegible] to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: BELONGS TO THE [illegible] FAMILY THAT GROUPS TOGETHER HSV-1 [illegible]
CC       [illegible] HCMV [illegible] EBV [illegible] AND VZV 4[illegible].
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; U43400; AAC54726.1; -.
SQ   SEQUENCE   439 AA;  51355 MW;  731214291E03[illegible]194 CRC64;

  Query Match             40.8%;  Score 42;  DB 1;  Length 439;
  Best Local Similarity   58.3%;  Pred. No. [illegible];
  Matches    7;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy        2 STCVLAGMATHP 13
            ::|||  ||:||
Db      325 NSCVLKAMASHP 336

RESULT  12
YGJI_ECOLI
[illegible]   YGJI_ECOLI     STANDARD;      PRT;   477 AA.
[illegible]   P42590;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   HYPOTHETICAL 52.1 KDA PROTEIN IN EBGC-UXAA INTERGENIC REGION.
GN   YGJI.
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC   Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12 / MG1655;
RX   MEDLINE=97426617; PubMed=9278503;
RA   Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA   Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA   Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA   Mau B., Shao Y.;
RT   "The complete genome sequence of Escherichia coli K-12.";
RL   Science 277:1453-1474(1997).
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE
CC       (POTENTIAL).
CC   -!- SIMILARITY: BELONGS TO THE XASA FAMILY OF TRANSPORTERS.
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; U18997; AAA57879.1; -.
DR   EMBL; AE000389; AAC76113.1; -.
DR   EcoGene; EG12720; ygjI.
DR   InterPro; IPR002027; -.
DR   Pfam; PF00324; aa_permeases; 1.
KW   Hypothetical protein; Transmembrane; Inner membrane; Transport.
FT   TRANSMEM      9     29       POTENTIAL.
FT   TRANSMEM     42     62       POTENTIAL.
FT   TRANSMEM     84    104       POTENTIAL.
FT   TRANSMEM    121    141       POTENTIAL.
FT   TRANSMEM    159    179       POTENTIAL.
FT   TRANSMEM    197    217       POTENTIAL.
FT   TRANSMEM    233    253       POTENTIAL.
FT   TRANSMEM    264    284       POTENTIAL.
FT   TRANSMEM    287    307       POTENTIAL.
FT   TRANSMEM    365    385       POTENTIAL.
FT   TRANSMEM    406    426       POTENTIAL.
FT   TRANSMEM    432    452       POTENTIAL.
SQ   SEQUENCE   477 AA;  52144 MW;  4395A[illegible]04E98[illegible]13 CRC64;

  Query Match             40.3%;  Score 41.5;  DB 1;  Length 477;
  Best Local Similarity   55.6%;  Pred. No. 2[illegible];
  Matches   10;  Conservative    2;  Mismatches    5;  Indels    1;  Gaps    1;

Qy        1 GSTCVLAGMATHPGEMPE 18
            ||   |  ||  :|| :||
Db      265 GSVQVFHGMAAYPG-LPE 281

RESULT  13
HIS6_RHOSH
ID   HIS6_RHOSH     STANDARD;      PRT;   252 AA.
AC   P50937;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   HISF PROTEIN (CYCLASE).
GN   HISF.
OS   Rhodobacter sphaeroides (Rhodopseudomonas sphaeroides).
OC   Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;
OC   Rhodobacter.
OX   NCBI_TaxID=1063;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN ATCC 17023 / 2.4.1 / NCIB 8253 / DSM 158;
RA   Oriol E.;
RL   Submitted (MAY-1995) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: CATALYZES THE CYCLIZATION REACTION THAT PRODUCES
CC       D-ERYTHRO-IMIDAZOLE-GLYCEROL PHOSPHATE.
CC   -!- PATHWAY: SIXTH STEP IN HISTIDINE BIOSYNTHETIC PATHWAY.
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE HISA / HISF FAMILY.
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; X87256; CAA60710.1; -.
DR   InterPro; IPR000570; -.
DR   Pfam; PF00977; His_biosynth; 1.
KW   Histidine biosynthesis.
SQ   SEQUENCE   252 AA;  26332 MW;  4E563AID[illegible]7DF07 CRC64;

  Query Match             39.8%;  Score 41;  DB 1;  Length 252;
  Best Local Similarity   53.3%;  Pred. No. [illegible];
  Matches    8;  Conservative    2;  Mismatches    5;  Indels    0;  Gaps    0;

Qy        1 GSTCVLAGMATHPGE 15
            |:: |||    ||||
Db      218 GASAVLAASIEHPGE 232

RESULT  14
Y4FP_RHISN
ID   Y4FP_RHISN     STANDARD;      PRT;   [illegible] AA.
AC   P55454;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   PROBABLE ABC TRANSPORTER PERIPLASMIC-BINDING PROTEIN Y4FP PRECURSOR.
GN   Y4FP.
OS   Rhizobium sp. (strain NGR234).
OG   Plasmid sym pNGR234a.
OC   Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC   Rhizobiaceae; Rhizobium.
```

```
OX   NCBI_TaxID [illegible];
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE 97[illegible]; PubMed [illegible];
RA   Freiberg C.A., [illegible] R., Broughton W.J., [illegible] A.,
RA   Perret X.;
RT   "Molecular basis of symbiosis between Rhizobium and legumes.";
RL   Nature 387:394-401(1997).
CC   -!- FUNCTION: PROBABLY PART OF A BINDING-PROTEIN-DEPENDENT TRANSPORT
CC       SYSTEM Y4EN/[illegible].
CC   -!- SUBCELLULAR LOCATION: PERIPLASMIC (POTENTIAL).
CC   -!- SIMILARITY: SOME [illegible] MARGINAL [illegible] [illegible]
CC       PROTEIN AND [illegible] [illegible].
CC   -!- CAUTION: IT IS UNCERTAIN WHETHER MET-[illegible] [illegible].
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AE000073; AAB9[illegible].1; -.
DR   Pfam; PF0154[illegible]; [illegible] bacterial_[illegible].
KW   Hypothetical protein; Transport; [illegible]; [illegible].
FT   SIGNAL        1     4[illegible]      [illegible] (POTENTIAL).
FT   CHAIN        4[illegible]   [illegible]      PROBABLE ABC TRANSPORTER [illegible]
FT                                  BINDING PROTEIN Y4[illegible].
SQ   SEQUENCE   [illegible] AA;  [illegible] MW;  [illegible] CRC64;

  Query Match             [illegible]%;  Score [illegible];  DB [illegible];  Length [illegible];
  Best Local Similarity   [illegible]%;  Pred. No. [illegible];
  Matches   [illegible];  Conservative   [illegible];  Mismatches   [illegible];  Indels   [illegible];  Gaps   [illegible];

Qy          [illegible] VLAGMA[illegible]EM [illegible]
               |    |||       :
Db      295 GSADVMAAYAA[illegible]QV[illegible] [illegible]


RESULT  15
[illegible]_CAEEL
ID   [illegible]_CAEEL      STANDARD;      PRT;   [illegible] AA.
AC   [illegible]01840;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   HYPOTHETICAL [illegible] PEPTIDE TRANSPORTER [illegible].
GN   F56F4.[illegible].
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   Murray [illegible], Wohldmann [illegible], Gillam [illegible].;
RL   Submitted (MAY-199[illegible]) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC   -!- SIMILARITY: BELONGS TO THE [illegible] FAMILY [illegible] TRANSPORTERS.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AF[illegible]; AAB[illegible].[illegible]; -.
DR   WormPep; F56F4.[illegible]; CE[illegible].
DR   InterPro; IPR00[illegible]; -.
```

```
DR   [illegible]
DR   [illegible]
DR   [illegible]
KW   Hypothetical protein; Transport; Transmembrane.
FT   TRANSMEM     [illegible]    4[illegible]       POTENTIAL.
FT   TRANSMEM     [illegible]    [illegible]       POTENTIAL.
FT   TRANSMEM     [illegible]    11[illegible]      POTENTIAL.
FT   TRANSMEM     [illegible]    1[illegible]      POTENTIAL.
FT   TRANSMEM     15[illegible]    17[illegible]      POTENTIAL.
FT   TRANSMEM     [illegible]    [illegible]      POTENTIAL.
FT   TRANSMEM     2[illegible]    2[illegible]      POTENTIAL.
FT   TRANSMEM     [illegible]    [illegible]      POTENTIAL.
FT   TRANSMEM     [illegible]    [illegible]      POTENTIAL.
FT   TRANSMEM     [illegible]    [illegible]      POTENTIAL.
FT   TRANSMEM     [illegible]    [illegible]      POTENTIAL.
FT   TRANSMEM     [illegible]    [illegible]      POTENTIAL.
SQ   SEQUENCE   [illegible] AA;  [illegible] MW;  A[illegible] CRC64;

  Query Match             [illegible]%;  Score [illegible];  DB [illegible];  Length [illegible];
  Best Local Similarity   4[illegible]%;  Pred. No. [illegible];
  Matches   [illegible];  Conservative   [illegible];  Mismatches   [illegible];  Indels   [illegible];  Gaps   [illegible];

Qy          [illegible] VLA [illegible] MA[illegible]ME[illegible]
                                 :
Db      4[illegible] [illegible]GSSWVLQARS[illegible]VGHF[illegible] [illegible]


Search completed: Aug[illegible] [illegible]
Job time: 4[illegible]
```

```
                     GenCore version 1.5
           Copyright (c) 1993 - 2001 Compugen Ltd.

OM protein - protein search, using sw model

Run on:          August 22, 2001, 14:27:24 ; Search time 112.12 Seconds
                                              (without alignments)
                                              22.121 Million cell updates/sec

Title:           US-08-659-254-5
Perfect score:   103
Sequence:        1 GSTCVLAGMATHFGEMPEQ 19

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters:      425026

  imum DB seq length: 0
  mum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       SPTREMBL_16:*
                1:  sp_archea:*
                2:  sp_bacteria:*
                3:  sp_fungi:*
                4:  sp_human:*
                5:  sp_invertebrate:*
                6:  sp_mammal:*
                7:  sp_mhc:*
                8:  sp_organelle:*
                9:  sp_phage:*
                10:  sp_plant:*
                11:  sp_rodent:*
                12:  sp_unclassified:*
                13:  sp_vertebrate:*
                14:  sp_virus:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                SUMMARIES
                 %
  lt            Query
  o.    Score   Match Length DB  ID                                  Description
----------------------------------------------------------------------------------------
     1     84    81.6    289  13  Q9DC08                         Q9dq08 xenopus lae
     2     77    74.8    281  13  Q9DC09                         Q9dq09 gallus gall
     3     66    64.1    285  6   Q9GML0                        Q9gml0 ovis aries
     4     66    64.1    287  13  Q9DEB4                         Q9deb4 oncorhynchu
     5     66    64.1    287  13  Q9DE06                         Q9de06 salvelinus
     6     65    63.1    285  13  Q9DC10                         Q9dq10 brachydanio
     7     45    43.7    634  2   O05402                        O05402 bacillus su
     8   44.5    43.2    446  2   Q9RSF8                        Q9rsf8 deinococcus
     9     44    42.7   1567  11  O35243                         O35243 mus muscul.
    10     44    42.7   1827  5   Q9U8B5                        Q9u8b5 drosophila
    11     44    42.7   1831  5   Q9W420                        Q9w420 drosophila
    12     44    42.7   1831  5   Q9U8B6                        Q9u8b6 drosophila
    13     43    41.7    275  4   Q9NY16                        Q9ny16 homo sapien
    14     43    41.7    459  5   Q26001                        Q26001 plasmodium
    15     43    41.7    793  4   Q9H9B0                        Q9h9b0 homo sapien
    16     43    41.7   1135  13  Q9YHD3                         Q9yhd3 xenopus lae
    17     43    41.7   1211  4   O94896                        O94896 homo sapien
    18     43    41.7   1259  5   O44971                        O44971 caenorhabdi
    19   42.5    41.3    482  13  Q918F7                         Q918f7 salmo salar
```

```
    20   42.5    41.3    521  13  Q918F6                        Q918f6 salmo salar
    21     42    40.8    386  5   Q19278                        Q19278 caenorhabdi
    22     42    40.8    435  14  O56261                         O56261 human herpe
    23     42    40.8    462  2   Q9RSD7                        Q9rsd7 deinococcus
    24     42    40.8    584  5   Q9VQE7                        Q9vqe7 drosophila
    25     42    40.8    891  4   Q9UHW6                        Q9uhw6 homo sapien
    26     42    40.8    891  11  Q9QY74                         Q9qy74 mus musculu
    27     42    40.8   1321  4   O75129                        O75129 homo sapien
    28     42    40.8   2138  5   Q9XZE3                        Q9xze3 amoeba prot
    29     42    40.8   4545  2   Q9X4W2                        Q9x4w2 vibrio chol
    30     42    40.8   4558  2   Q9KS12                        Q9ks12 vibrio chol
    31   41.5    40.3    214  3   Q9P316                        Q9p316 neurospora
    32   41.5    40.3    247  5   Q20908                        Q20908 caenorhabdi
    33     41    39.8    124  5   Q9NHG6                        Q9nhg6 plasmodium
    34     41    39.8    205  4   Q15918                        Q15918 homo sapien
    35     41    39.8    210  2   Q9PEX0                        Q9pex0 xylella fas
    36     41    39.8    236  1   O27417                        O27417 methanobact
    37     41    39.8    269  2   Q9RY29                        Q9ry29 deinococcus
    38     41    39.8    274  1   Q9V118                        Q9v118 pyrococcus
    39     41    39.8    328  2   Q9PP26                        Q9pp26 campylobact
    40     41    39.8    421  11  Q62515                         Q62515 mus musculu
    41     41    39.8    429  3   O74802                        O74802 schizosacch
    42     41    39.8    445  3   O94214                        O94214 mortierella
    43     41    39.8    445  3   O94215                        O94215 mortierella
    44     41    39.8    445  3   Q9UVW5                        Q9uvw5 mortierella
    45     41    39.8    803  3   O74402                        O74402 schizosacch

                                ALIGNMENTS

RESULT    1
Q9DG08
ID   Q9DG08      PRELIMINARY;      PRT;    289 AA.
AC   Q9DG08;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN (FRAGMENT).
GN   STAR.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=OVARY;
RX   PubMed=11064158;
RA   Bauer M.P., Bringham J.T., Langenau D.M., Johnson A.L., Goetz F.W.;
RT   "Conservation of steroidogenic acute regulatory (StAR) protein
RT   structure and expression in vertebrates.";
RL   Mol. Cell. Endocrinol. 168:119-125(2000).
DR   EMBL; AF220437; AAG28595.1; -.
FT   NON_TER        1       1
FT   NON_TER      289     289
SQ   SEQUENCE   289 AA;  32066 MW;  C77...  2AB*86BD8E  7FC64;

  Query Match             81.6%;  Score 84;  DB 13;  Length 289;
  Best Local Similarity   78.9%;  Pred. No. 5.8e-06;
  Matches   15;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GSTCVLAGMATHFGEMPEQ 19
              ||||:||||:| || ||||
Db        198 GSTCILAGMSIRFGGMPEQ 216


RESULT    2
Q9DG09
ID   Q9DG09      PRELIMINARY;      PRT;    281 AA.
AC   Q9DG09;
DT   01 MAR 2001 (TrEMBLrel. 16, Created)
DT   01 MAR 2001 (TrEMBLrel. 16, Last sequence update)
DT   01 MAR 2001 (TrEMBLrel. 16, Last annotation update)
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN.
GN   STAR.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; [illegible]
OC   Archosauria; Aves; Neognathae; [illegible]
OC   Gallus.
OX   NCBI_TaxID=[illegible];
RN   [1]
RP   SEQUENCE FROM N.A.
RX   PubMed=1106[illegible];
RA   Bauer M.P., Bridgham J.T., Langenau D.M., [illegible] W.
RT   "Conservation of steroidogenic acute regulatory (StAR) protein
RT   structure and expression in vertebrates.";
RL   Mol. Cell. Endocrinol. [illegible]
DR   EMBL; AF2204[illegible]; AAG[illegible]
SQ   SEQUENCE   281 AA;  [illegible] MW;  [illegible]

  Query Match             [illegible];  Score [illegible];  DB [illegible];  Length [illegible];
  Best Local Similarity   78.[illegible];  Pred. No. [illegible];
  Matches   15;  Conservative   [illegible];  Mismatches   [illegible];  Indels   [illegible];  Gaps   [illegible];

Qy          1 GSICVLA[illegible]MA[illegible]HGEM[illegible]EQ 1[illegible]
              [illegible]
Db        194 GSICVLAGMS[illegible]HGAM[illegible]EQ 2[illegible]2


RESULT 3
Q9GMD0
ID   Q9GMD0      PRELIMINARY;      PRT;   2[illegible] AA.
AC   Q9GMD0;
DT   01 MAR 2001 (TrEMBLrel. 16, Created)
DT   01 MAR 2001 (TrEMBLrel. 16, Last sequence update)
DT   01 MAR 2001 (TrEMBLrel. 16, Last annotation update)
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN.
GN   STAR.
OS   Ovis aries (Sheep).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; [illegible]
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; [illegible]
OC   Bovidae; Caprinae; Ovis.
OX   NCBI_TaxID=9940;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   West L.A., Horvat R.J., [illegible] D.A., [illegible]
RA   Niswender G.D.;
RT   "Steroidogenic Acute Regulatory Protein and Peripheral [illegible]
RT   Benzodiazepine Receptor Associate at the Mitochondrial Membrane.";
RL   Submitted (JUN-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF290202; AAG[illegible]
SQ   SEQUENCE   28[illegible] AA;  [illegible] MW;  AB[illegible]

  Query Match             [illegible];  Score [illegible];  DB [illegible];  Length [illegible];
  Best Local Similarity   [illegible];  Pred. No. [illegible];
  Matches   13;  Conservative   2;  Mismatches   [illegible];  Indels   [illegible];  Gaps   [illegible];

Qy          1 GSICVLAGMAI[illegible]GEM[illegible] 1[illegible]
              [illegible]
Db        194 GSMCVLAG[illegible]EM[illegible]


RESULT 4
Q9DEB4
ID   Q9DEB4      PRELIMINARY;      PRT;   28[illegible] AA.
AC   Q9DEB4;
DT   01-MAR 2001 (TrEMBLrel. 16, Created)
DT   01 MAR 2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR 2001 (TrEMBLrel. 16, Last annotation update)
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN.
GN   STAR.
OS   Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC   Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX   NCBI_TaxID=[illegible];
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=HEAD KIDNEY;
RA   [illegible] Kusakabe M., [illegible] Young G.;
RT   "Molecular [illegible] steroidogenic acute regulatory (StAR) protein
RT   cDNA from rainbow trout.";
RL   Submitted ([illegible]) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB[illegible]
SQ   SEQUENCE   28[illegible] AA;  [illegible] MW;  [illegible]

  Query Match             [illegible];  Score [illegible];  DB [illegible];  Length 28[illegible];
  Best Local Similarity   68.[illegible];  Pred. No. [illegible];
  Matches   [illegible];  Conservative   [illegible];  Mismatches   [illegible];  Indels   [illegible];  Gaps   [illegible];

Qy          1 [illegible]VLAGMA[illegible]HGEM[illegible]EQ 1[illegible]
              [illegible]
Db        [illegible] [illegible]TLAGMS[illegible]EMLEQ 2[illegible]


RESULT 5
Q9[illegible]
ID   Q9[illegible]      PRELIMINARY;      PRT;   2[illegible] AA.
AC   Q9[illegible];
DT   01 MAR 2001 (TrEMBLrel. 16, Created)
DT   01 MAR 2001 (TrEMBLrel. 16, Last sequence update)
DT   01 MAR 2001 (TrEMBLrel. 16, Last annotation update)
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN.
GN   STAR.
OS   Gasterosteus [illegible] (Three-spined stickleback).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC   [illegible]; Gasterosteiformes; Gasterosteidae; Gasterosteus.
OX   NCBI_TaxID=[illegible];
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=BRAIN, [illegible], HEAD KIDNEY, SKIN, AND LIVER;
RA   [illegible];
RT   "Characterization of the [illegible] StAR and MLN64 [illegible].";
RL   Submitted ([illegible]) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF[illegible]; AA[illegible]
SQ   SEQUENCE   2[illegible] AA;  [illegible] MW;  [illegible]

  Query Match             [illegible];  Score [illegible];  DB [illegible];  Length 28[illegible];
  Best Local Similarity   [illegible];  Pred. No. [illegible];
  Matches   [illegible];  Conservative   [illegible];  Mismatches   [illegible];  Indels   [illegible];  Gaps   [illegible];

Qy          1 [illegible]VLAGMAI[illegible]MLEQ 1[illegible]
              [illegible]
Db        [illegible] [illegible]AGM[illegible]MLEQ [illegible]


RESULT 6
Q9DE[illegible]
ID   Q9[illegible]      PRELIMINARY;      PRT;   2[illegible] AA.
AC   Q9[illegible];
DT   [illegible] MAR [illegible] (TrEMBLrel. [illegible], Created)
DT   [illegible] MAR [illegible] (TrEMBLrel. [illegible], Last sequence update)
DT   [illegible] MAR [illegible] (TrEMBLrel. [illegible], Last annotation update)
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN.
GN   STAR.
OS   [illegible] (Zebrafish) (Brachydanio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC   Cypriniformes; Cyprinidae; Rasborinae; Danio.
OX   NCBI_TaxID=[illegible];
RN   [illegible]
RP   SEQUENCE FROM N.A.
```

```
RX   PubMed=11064158;
RA   Bauer M.P., Bridgham J.T., Langenau D.M., Johnson A.L., Goetz F.W.;
RT   "Conservation of steroidogenic acute regulatory (StAR) protein
RT   structure and expression in vertebrates.";
RL   Mol. Cell. Endocrinol. 168:119-125(2000)
DR   EMBL; AF220435; AAG28593.1; -.
SQ   SEQUENCE   285 AA;  31759 MW;  3F7251C07 716078 CRC64;

  Query Match             63.1%;  Score 63;  DB 1;  Length 285;
  Best Local Similarity   68.4%;  Pred. No. 0.0078;
  Matches   13;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 GSTCVLAGMATHFGEMPEQ 19
              |||| ||||:|    ||||
Db        192 GSTCFLAGMSTQHPGMPEQ 210


RESULT   7
O05402
ID   O05402      PRELIMINARY;      PRT;   634 AA.
AC   O05402;
     01-JUL-1997 (TrEMBLrel. 04, Created)
     01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   HYPOTHETICAL 72.2 KDA PROTEIN.
GN   YRHL.
OS   Bacillus subtilis.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group;
OC   Bacillus/Staphylococcus group; Bacillus.
OX   NCBI_TaxID=1423;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=168;
RA   Sorokin A., Bolotin A., Purnelle B., Hilbert H., Lauber J.,
RA   Duesterhoeft A., Ehrlich S.D.;
RL   Submitted (APR-1997) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=168;
RX   MEDLINE=98044033; PubMed=9384377;
RA   Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA   Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA   Borriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA   Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA   Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA   Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
RA   Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA   Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA   Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
     Guiseppi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,
     Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
R    Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA   Kobayashi Y., Koetter P., Koningstein G., Krogh S., Kumano M.,
RA   Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA   Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
RA   Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA   Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA   Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA   Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA   Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
RA   Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA   Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
RA   Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA   Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Togoni A.,
RA   Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA   Viari A., Wambutt R., Wedler E., Wedler H., Weitzenegger T.,
RA   Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA   Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
RT   "The complete genome sequence of the gram positive bacterium Bacillus
RT   subtilis.";
RL   Nature 390:249-256(1997).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=168;
RA   Kunst F., Ogasawara N., Yoshikawa H., Danchin A.;
RL   Submitted (NOV-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U93874; AAB80869.1; -.
DR   EMBL; Z99117; CAB14656.1; -.
DR   InterPro; IPR002656; -.
DR   Pfam; PF01757; DUF33; 2.
KW   Hypothetical protein.
SQ   SEQUENCE   634 AA;  72218 MW;  1F7AF4A9742 20B CRC64;

  Query Match             43.7%;  Score 43;  DB 2;  Length 634;
  Best Local Similarity   53.3%;  Pred. No. 9;
  Matches    8;  Conservative    3;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          3 TCVLAGMATHFGEMP 17
              ||:|| ::  | | |
Db        341 TCILAELSYHFLEKP 355


RESULT   8
Q9RSE8
ID   Q9RSE8      PRELIMINARY;      PRT;   446 AA.
AC   Q9RSE8;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   CONSERVED HYPOTHETICAL PROTEIN.
GN   DR2176.
OS   Deinococcus radiodurans.
OC   Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.
OX   NCBI_TaxID=1299;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=R1;
RX   MEDLINE=20036896; PubMed=10567266;
RA   White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA   Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA   Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA   Vamathevan J.J., Lam P., McDonald L., Utterback T., Zalewski C.,
RA   Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA   Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA   Fraser C.M.;
RT   "Genome sequence of the radioresistant bacterium Deinococcus
RT   radiodurans R1.";
RL   Science 286:1571-1577(1999).
DR   EMBL; AE002051; AAF11725.1; -.
DR   TIGR; DR2176; -.
DR   InterPro; IPR000644; -.
DR   InterPro; IPR002550; -.
DR   Pfam; PF00571; CBS; 2.
DR   Pfam; PF01595; DUF21; 1.
DR   SMART; SM00116; CBS; 1.
SQ   SEQUENCE   446 AA;  49355 MW;  7B82CF197 2CABE4 CRC64;

  Query Match             43.2%;  Score 44.5;  DB 2;  Length 446;
  Best Local Similarity   64.3%;  Pred. No. 30;
  Matches    9;  Conservative    3;  Mismatches    1;  Indels    1;  Gaps    1;

Qy          6 LAG-MATHFGEMPE 18
              ||| | :|:|::||
Db        388 LACFMTSHFGDIPE 401


RESULT   9
O35243
ID   O35243      PRELIMINARY;      PRT;   1567 AA.
AC   O35243;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
```

```
DE   ANTIGEN CONTAINING [illegible]
DE   (FRAGMENT).
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; [illegible]
OC   Mammalia; Eutheria; Rodentia; [illegible]
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=9611[illegible]; PubMed=[illegible];
RA   Benayahu D., [illegible] Wientroub S.;
RT   "Monoclonal antibodies [illegible]";
RL   J. Bone Miner. Res. [illegible]
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Fisher L.W., Mercelli A., Benayahu D.;
RL   Submitted ([illegible]) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF01[illegible]; AAB6[illegible]
DR   InterPro; IPR00[illegible];
DR   Pfam; PF0217[illegible]
DR   SMART; SM00[illegible]
FT   NON_TER      1      1
SQ   SEQUENCE   [illegible] AA;  [illegible] MW;  [illegible] CRC64;

  Query Match              [illegible]%;  Score 11;  [illegible]
  Best Local Similarity    [illegible]%;  Pred. No. [illegible];
  Matches    [illegible];  Conservative  [illegible];  Mismatches  [illegible]

Qy          1 [illegible]VLA[illegible]MA[illegible]
              [illegible]
Db        17[illegible] GSACIVE[illegible]


RESULT 10
Q9DBB5
ID   Q9DBB5     PRELIMINARY;   PRT;   [illegible] AA.
AC   Q9DBB5;
DT   01 MAY 2000 (TrEMBLrel. [illegible], Created)
DT   01 MAY 2000 (TrEMBLrel. [illegible], Last sequence update)
DT   01 MAR 2001 (TrEMBLrel. [illegible], Last annotation update)
DE   SPI6 PROTEIN.
GN   SPI6 OR CG12[illegible].
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Chiang P.W.;
RT   "Characterization of Drosophila [illegible]";
RL   Submitted (NOV-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF104401; AAF14[illegible]
DR   FlyBase; FBgn00[illegible]; Spi6.
DR   InterPro; IPR00[illegible];
DR   InterPro; IPR00[illegible];
DR   Pfam; PF00575; S1; 1.
DR   SMART; SM00[illegible];
SQ   SEQUENCE   [illegible] AA;  [illegible] MW;  [illegible] CRC64;

  Query Match              [illegible]%;  Score 11;  [illegible]
  Best Local Similarity    [illegible]%;  Pred. No. [illegible];
  Matches    [illegible];  Conservative  [illegible];  Mismatches  [illegible]

Qy          [illegible] LAGMAIHE[illegible]
              [illegible]
Db        533 IGGLAKHE[illegible]


RESULT 11
Q9W420
ID   Q9W420     PRELIMINARY;   PRT;   [illegible] AA.
```

```
AC   Q9W4[illegible];
DT   [illegible] MAY [illegible] (TrEMBLrel. [illegible], Created)
DT   [illegible] MAY [illegible] (TrEMBLrel. [illegible], Last sequence update)
DT   [illegible] MAR [illegible] (TrEMBLrel. [illegible], Last annotation update)
DE   [illegible]PROTEIN.
GN   [illegible]
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RX   MEDLINE=[illegible]; PubMed=[illegible];
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
DR   EMBL; AE00[illegible]437; AAF[illegible]
DR   FlyBase; FBgn00[illegible]
DR   InterPro; IPR00[illegible]
DR   InterPro; IPR[illegible]
DR   Pfam; PF[illegible]
DR   SMART; SM[illegible]
SQ   SEQUENCE   [illegible] AA;  [illegible] MW;  [illegible]441[illegible] CRC64;

  Query Match              [illegible]%;  Score [illegible];  DB [illegible];  Length [illegible];
  Best Local Similarity    [illegible]%;  Pred. No. [illegible];
  Matches    [illegible];  Conservative  [illegible];  Mismatches  [illegible];  Indels  [illegible];  Gaps  [illegible];

Qy          [illegible] LAGMA[illegible]

Db        [illegible] [illegible]EAK[illegible]


RESULT 1[illegible]
[illegible]
```

```
ID   Q9U8B6      PRELIMINARY;      PRT;  1831 AA.
AC   Q9U8B6;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   SPT6 PROTEIN.
GN   SPT6 OR CG12225.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID 7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Chiang P.W.;
RT   "Characterization of Drosophila SPT6 homolog.";
RL   Submitted (NOV-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF104400; AAF14114.1; -.
DR   FlyBase; FBgn0028982; spt6.
DR   InterPro; IPR000980; -.
DR   InterPro; IPR003029; -.
DR   Pfam; PF00575; S1; 1.
     SMART; SM00252; SH2; 1.
     SEQUENCE   1831 AA;  208597 MW;  877AB[illegible]45CAB180 CRC64;

  Query Match             42.7%;  Score 44;  DB 5;  Length 1831;
  Best Local Similarity   57.1%;  Pred. No. 1.6e+02;
  Matches    8;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          6 LAGMAIHFGEMPEQ 19
                :| | |||  |||
Db        537 ICGFAKHFGLTPEQ 550


RESULT  13
Q9NY16
ID   Q9NY16      PRELIMINARY;      PRT;  275 AA.
AC   Q9NY16;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   MESENCHYMAL STEM CELL PROTEIN DSC43.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID 9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   van den Bos C., Mbalaviele G., Thiede M.;
     "cDNA DSC43 expressed by osteogenic human mesenchymal stem cells.";
     Submitted (MAR 2000) to the EMBL/GenBank/DDBJ databases.
     -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY
DR   EMBL; AF242768; AAF65445.1; -.
DR   InterPro; IPR000582; -.
DR   InterPro; IPR000822; -.
DR   Pfam; PF00096; zf-C2H2; 8.
DR   PRINTS; PR00048; ZINCFINGER.
DR   ProDom; PD002965; -; 1.
DR   PROSITE; PS00028; ZINC_FINGER_C2H2; 8.
KW   DNA-binding; Metal-binding; Nuclear protein; Zinc finger.
SQ   SEQUENCE   275 AA;  30781 MW;  52BF7[illegible]0BE466D CRC64;

  Query Match             41.7%;  Score 43;  DB 4;  Length 275;
  Best Local Similarity   47.1%;  Pred. No. 42;
  Matches    8;  Conservative    2;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          1 GSICVLAGMATHFGEMP 17
              ||:|:     || || |
Db        144 GSSCLARHRRTHTGERP 160
```

```
RESULT  14
Q26001
ID   Q26001      PRELIMINARY;      PRT;  459 AA.
AC   Q26001;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   RABGDI PROTEIN.
GN   RABGDI.
OS   Plasmodium falciparum (isolate 3D7).
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID 36329;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=3D7;
RX   MEDLINE 97[illegible]33; PubMed=884467[illegible];
RA   Attal G., Langsley G.;
RT   "A Plasmodium falciparum homologue of a rab specific GDP dissociation
RT   inhibitor.";
RL   Mol. Biochem. Parasitol. 79:91-96(1996).
DR   EMBL; X93166; CAA63653.1; -.
DR   HSSP; P21856; 1GND.
DR   InterPro; IPR002005; -.
DR   Pfam; PF00996; GDI; 1.
DR   PRINTS; PR00891; RABGDIREP.
SQ   SEQUENCE   459 AA;  52311 MW;  [illegible]A47B[illegible]BEA CRC64;

  Query Match             41.7%;  Score 43;  DB 5;  Length 459;
  Best Local Similarity   41.7%;  Pred. No. 55;
  Matches    5;  Conservative    6;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 CVLAGMAIHFGE 15
              |:|:|:  :|:
Db         17 CILSGLLSHYGK 28


RESULT  15
Q9H9B0
ID   Q9H9B0      PRELIMINARY;      PRT;  793 AA.
AC   Q9H9B0;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   CDNA FLJ12883 FIS, CLONE NT2RP200[illegible], WEAKLY SIMILAR TO VACUOLAR
DE   PROTEIN SORTING-ASSOCIATED PROTEIN VPS8.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID 9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA   Nishikawa T., Nagai K., Sugano S., Takahashi-Fujii A., Hara H.,
RA   Tanase T., Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K.,
RA   Arita M., Nabekura T., Ishii S., Kawai Y., Saito K., Yamamoto J.,
RA   Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y., Isshima A.;
RT   "NEDO human cDNA sequencing project.";
RL   Submitted (AUG 2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AK022945; BAB14322.1; -.
SQ   SEQUENCE   793 AA;  91227 MW;  5E[illegible]A[illegible]D64 CRC64;

  Query Match             41.7%;  Score 43;  DB 4;  Length 793;
  Best Local Similarity   50.0%;  Pred. No. 42;
  Matches    8;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          3 TCVLAGMAIHFGEMPE 18
               :| ||| |   |::||
Db         95 SCCLAGRAYPIGDIPE 110
```

```
Search completed: August [illegible]
Job time: [illegible] sec
```

```
                         GenCore version 4.5
              Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on:          August 22, 2001, 14:27:22 ; Search time 112.12 Seconds
                                              (without alignments)
                                              14.160 Million cell updates/sec

Title:           US-08-659-254-4
Perfect score:   66
Sequence:        1 ALGILNNQEGWK 12

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters:       425026

 mum DB seq length: 0
 mum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       SPTREMBL_16:*
                1:  sp_archea:*
                2:  sp_bacteria:*
                3:  sp_fungi:*
                4:  sp_human:*
                5:  sp_invertebrate:*
                6:  sp_mammal:*
                7:  sp_mhc:*
                8:  sp_organelle:*
                9:  sp_phage:*
                10:  sp_plant:*
                11:  sp_rodent:*
                12:  sp_unclassified:*
                13:  sp_vertebrate:*
                14:  sp_virus:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                              SUMMARIES
              %
 lt.        Query
No.  Score  Match Length DB  ID                           Description
-----------------------------------------------------------------------------
   1    55   83.3    285  6  Q9GMD0                       Q9gmd0 ovis aries
   2    48   72.7    783  2  Q45554                       Q45554 bacillus sp
   3    48   72.7    824  2  Q9F2I6                       Q9f2i6 bacillus sp
   4    46   69.7    281 13   Q9DG09                      Q9dg09 gallus gall
   5    43   65.2    289 13   Q9DC98                      Q9dc98 xenopus lae
   6    43   65.2    821  2  Q59241                       Q59241 bacillus sp
   7    41   62.1    175 11   Q9QUS9                      Q9qus9 mus muscul
   8    41   62.1    175 11   Q9JMH1                      Q9jmh1 mus musculu
   9    41   62.1    905 14   Q67483                      Q67483 banzi virus
  10    41   62.1   1815  5  Q22541                       Q22541 caenorhabdi
  11  40.5   61.4    329 14   Q9EIV8                      Q9eiv8 edge hill v
  12  40.5   61.4    329 14   Q9EIV5                      Q9eiv5 edge hill v
  13    40   60.6    331  2  Q9F3B9                       Q9f3b9 streptomyce
  14    40   60.6    695  8  Q31773                       Q31773 atropa bell
  15    40   60.6   1049  5  O00851                       O00851 entamoeba h
  16    40   60.6   1463  2  O86919                       O86919 staphylococ
  17    39   59.1    132  2  Q9P9Q8                       Q9p9q8 xylella fas
  18    39   59.1    200  2  Q9JV98                       Q9jv98 neisseria m
  19    39   59.1    260  8  O78318                       O78318 arabidopsis
```

```
  20    39   59.1    346 10   O24417                      O24417 cucurbita m
  21    39   59.1    349 13   O42258                      O42258 xenopus lae
  22    39   59.1    505  2  Q9K079                       Q9k079 neisseria m
  23    39   59.1   1445  5  O01817                       O01817 caenorhabdi
  24    39   59.1   1711  2  P96311                       P96311 [illegible]
  25    38   57.6    102  4  Q9H2I4                       Q9h2i4 homo sapien
  26    38   57.6    171  2  Q9KU46                       Q9ku46 vibrio chol
  27    38   57.6    181  2  Q9I341                       Q9i341 pseudomonas
  28    38   57.6    311  2  Q51863                       Q51863 tetragenoco
  29    38   57.6    330  5  O76238                       O76238 tethya aura
  30    38   57.6    372  2  O05804                       O05804 mycobacteri
  31    38   57.6    388  2  O53111                       O53111 mycoplasma
  32    38   57.6   1289  2  Q9WZ87                       Q9wz87 thermotoga
  33    38   57.6   1943  2  O25331                       O25331 helicobacte
  34    38   57.6   2293 14   Q83422                      Q83422 [illegible]
  35  37.5   56.8    329 14   Q9EIV7                      Q9eiv7 edge hill v
  36  37.5   56.8    329 14   Q9EIV6                      Q9eiv6 edge hill v
  37    37   56.1     66  2  Q9FCL9                       Q9fcl9 streptomyce
  38    37   56.1     99  2  O68792                       O68792 yersinia pe
  39    37   56.1    105  2  Q9RI19                       Q9ri19 yersinia pe
  40    37   56.1    126 10   O24240                      O24240 prunus arme
  41    37   56.1    285 13   Q9DEI0                      Q9dei0 brachydanio
  42    37   56.1    362  4  O95884                       O95884 homo sapien
  43    37   56.1    369  5  O96622                       O96622 dictyosteli
  44    37   56.1    393 11   O08719                      O08719 rattus norv
  45    37   56.1    393 11   P70429                      P70429 mus musculu

                              ALIGNMENTS

RESULT    1
Q9GMD0
ID   Q9GMD0      PRELIMINARY;      PRT;    285 AA.
AC   Q9GMD0;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN.
GN   STAR.
OS   Ovis aries (Sheep).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Caprinae; Ovis.
OX   NCBI_TaxID=9940;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   West L.A., Horvat R.D., Roess D.A., Barisas B.G., Juengel J.L.,
RA   Niswender G.D.;
RT   "Steroidogenic Acute Regulatory Protein and Peripheral-Type
RT   Benzodiazepine Receptor Associate at the Mitochondrial Membrane.";
RL   Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF290202; AAG02464.1; -.
SQ   SEQUENCE   285 AA;  31895 MW;  [illegible];

  Query Match              83.3%;  Score 55;  DB 6;  Length 285;
  Best Local Similarity    83.3%;  Pred. No. [illegible];
  Matches   10;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        1 ALGILNNQEGWK 12
            |||||: |||||
Db       86 ALGILKDQEGWK 97


RESULT    2
Q45554
ID   Q45554      PRELIMINARY;      PRT;    783 AA.
AC   Q45554;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   ENDOGLUCANASE PRECURSOR (EC 3.2.1.4) (ENDO-1,4-BETA-GLUCANASE)
DE   (ALKALINE [illegible]ASE
OS   Bacillus sp.
OC   Bacteria; Firmicutes; [illegible]
OC   Bacillus/Staph[illegible]
OX   NCBI_TaxID=140[illegible]
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN [illegible]
RA   Miyatake M., [illegible]
RL   Submitted (MAY [illegible]
CC   -!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS [illegible]
CC       LINKAGES IN CELLULOSE (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO [illegible] FAMILY [illegible]
CC       HYDROLASES)
DR   EMBL; D8[illegible]2[illegible]; BAA[illegible]
DR   HSSP; [illegible]5465; 1A3H.
DR   InterPro; IPR00[illegible]
DR   Pfam; PF001[illegible]; Cellulase; 1.
DR   PROSITE; PS[illegible]; GLYCOSYL HYDROL[illegible] NEW[illegible]
KW   Cellulose degradation; Hydrolase; [illegible]
FT   SIGNAL       [illegible]        POTENTIAL.
FT   CHAIN        [illegible]        ALKALINE [illegible]
FT   ACT_SITE     [illegible]        PROTON DONOR (BY SIMILARITY).
FT   ACT_SITE     [illegible]        NUCLEOPHILE (BY SIMILARITY).
SQ   SEQUENCE   [illegible] AA;  [illegible] MW;  [illegible]

  Query Match   [illegible]
  Best Local Similarity   [illegible]
  Matches   [illegible];  Conservative   [illegible];  Mismatches   [illegible]

Qy          3 GIENNQEGWK 12
                  |:
Db        200 GIENNEEGWK [illegible]


RESULT 3
Q9F216
ID   Q9F216      PRELIMINARY;      PRT;   [illegible] AA.
AC   Q9F216;
DT   01-MAR-2001 (TrEMBLrel. [illegible], Created)
DT   01-MAR-2001 (TrEMBLrel. [illegible], Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. [illegible], Last annotation update)
DE   CELLULASE.
OS   Bacillus sp.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group;
OC   Bacillus/Staphylococcus group; Bacillus.
OX   NCBI_TaxID=1409;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=KSM-S[illegible];
RA   Hakamada Y., Hatada Y., Koike K., Yoshimatsu T., [illegible]
RA   Kobayashi T., Ito S.;
RT   "Deduced amino acid sequence [illegible]
RT   thermostable, alkaline cellulase from an alkaliphilic Bacillus
RT   strain.";
RL   Biosci. Biotechnol. Biochem. 64:[illegible]
DR   EMBL; AB[illegible]; BAA[illegible]
SQ   SEQUENCE   824 AA;  [illegible] MW;  [illegible]

  Query Match   [illegible]
  Best Local Similarity   [illegible]
  Matches   [illegible];  Conservative   [illegible];  Mismatches   [illegible]

Qy          3 GIENNQEGWK 12
                 || |:|||
Db        200 GIENNEEGWK [illegible]


RESULT 4
Q9DG09
ID   Q9DG09      PRELIMINARY;      PRT;   [illegible] AA.
AC   [illegible]
DT   [illegible] MAR [illegible] (TrEMBLrel. [illegible], Created)
DT   [illegible] MAR [illegible] (TrEMBLrel. [illegible], Last sequence update)
DT   [illegible] MAR [illegible] (TrEMBLrel. [illegible], Last annotation update)
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN.
GN   STAR.
OS   [illegible]
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   [illegible]
OC   [illegible]
OX   NCBI_TaxID=[illegible]
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=[illegible]
RA   [illegible]
RT   "[illegible] (StAR) protein
RT   structure and expression in vertebrates.";
RL   Mol. [illegible]
DR   EMBL; AF[illegible]; AAG[illegible]
SQ   SEQUENCE   [illegible] AA;  [illegible] MW;  [illegible]

  Query Match   [illegible]
  Best Local Similarity   [illegible]
  Matches   [illegible];  Conservative   [illegible];  Mismatches   [illegible]

Qy          [illegible] NNQEGWK [illegible]
                 [illegible]
Db          [illegible] GWK [illegible]


RESULT 5
[illegible]
ID   [illegible]      PRELIMINARY;      PRT;   [illegible] AA.
AC   [illegible]
DT   [illegible] MAR [illegible] (TrEMBLrel. [illegible], Created)
DT   [illegible] MAR [illegible] (TrEMBLrel. [illegible], Last sequence update)
DT   [illegible] MAR [illegible] (TrEMBLrel. [illegible], Last annotation update)
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN (FRAGMENT).
GN   STAR.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=OVARY;
RX   MEDLINE=[illegible]
RA   Bauer M.P., Bridgham J.T., Langenau D.M., [illegible] A.L., [illegible]
RT   "Conservation of steroidogenic acute regulatory (StAR) protein
RT   structure and expression in vertebrates.";
RL   Mol. Cell. Endocrinol. [illegible]
DR   EMBL; AF[illegible]; AAG[illegible]
FT   NON_TER   [illegible]
FT   NON_TER   [illegible]
SQ   SEQUENCE   [illegible] AA;  [illegible] MW;  [illegible]

  Query Match   [illegible]
  Best Local Similarity   [illegible]
  Matches   [illegible];  Conservative   [illegible];  Mismatches   [illegible]

Qy          [illegible] NNQEGWK [illegible]

Db          [illegible] GWK [illegible]


RESULT 6
[illegible]
ID   [illegible]      PRELIMINARY;      PRT;   [illegible] AA.
AC   [illegible]
DT   [illegible] MAR [illegible] (TrEMBLrel. [illegible], Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   ENDO-1,4-BETA-GLUCANASE (EC 3.2.1.4) (CELLULASE) (ENDOGLUCANASE)
DE   (CARBOXYMETHYL CELLULASE).
OS   Bacillus sp.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group;
OC   Bacillus/Staphylococcus group; Bacillus.
OX   NCBI_TaxID=1409;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=KSM-64;
RA   Sumitomo N., Ozaki K., Ito S.;
RL   Submitted (MAY-1993) to the EMBL/GenBank/DDBJ databases.
CC   -!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF 1,4-BETA-D-GLUCOSIDIC
CC       LINKAGES IN CELLULOSE.
DR   EMBL; M84963; AAA73189.1; -.
DR   HSSP; O85465; 1A3H.
DR   InterPro; IPR001547; -.
DR   Pfam; PF00150; cellulase; 1.
DR   PROSITE; PS00659; GLYCOSYL_HYDROL_F5; UNKNOWN_1.
KW   Hydrolase; Glycosidase.
     SEQUENCE   821 AA;  90910 MW;  7354[illegible] CRC64;

  Query Match             65.2%;  Score 43;  DB 2;  Length 821;
  Best Local Similarity   77.8%;  Pred. No. [illegible];
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    [illegible];  Gaps    0;

Qy          3 GILNNQEGW 11
              || ||:|||
Db        199 GIPNNEEGW 207


RESULT   7
Q9QUS9
ID   Q9QUS9      PRELIMINARY;      PRT;   175 AA.
AC   Q9QUS9;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   REG III DELTA PRECURSOR.
GN   REG3D.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090,
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Abe M., Nata K., Tomioka-Kumagai T., [illegible] S., Ito S., Okamoto H.;
RT   "Identification of a novel Reg family gene, Reg III delta, and mapping
     of all three types of Reg family genes in 75 kilobase mouse genomic
     region.";
     Submitted (NOV-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB035204; BAA88564.1; -.
DR   EMBL; AB035205; BAA88565.1; -.
DR   HSSP; P05451; 1LIT.
DR   InterPro; IPR001304; -.
DR   Pfam; PF00059; lectin_c; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
DR   SMART; SM00034; CLECT; 1.
KW   Signal.
FT   SIGNAL        1     26       POTENTIAL.
SQ   SEQUENCE   175 AA;  20037 MW;  6A88[illegible]12B76B53 CRC64;

  Query Match             62.1%;  Score 41;  DB 11;  Length 175;
  Best Local Similarity   58.3%;  Pred. No. 9.1;
  Matches    7;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 ALGILNNQEGWK 12
              :|| | |: |||
Db        110 SLGSLPNENGWK 121
```

```
RESULT   8
Q9JMH1
ID   Q9JMH1      PRELIMINARY;      PRT;   175 AA.
AC   Q9JMH1;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   INGAP RELATED PROTEIN.
GN   INGAPRP.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BDF 1; TISSUE=DUODENUM;
RX   MEDLINE=20033449; PubMed=10564727;
RA   Sasahara K., Yamaoka T., Moritani M., [illegible] K., Kuroda Y.,
RA   Itakura M.;
RT   "Molecular cloning and tissue-specific expression of a new member of
RT   the regenerating protein family, islet neogenesis-associated protein-
RT   related protein.";
RL   Biochim. Biophys. Acta 1500:142-149(2000).
DR   EMBL; AB028625; BAA92141.1; -.
DR   InterPro; IPR001304; -.
DR   Pfam; PF00059; lectin_c; 1.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
DR   SMART; SM00034; CLECT; 1.
SQ   SEQUENCE   175 AA;  20021 MW;  [illegible]112B76B53 [illegible]4;

  Query Match             62.1%;  Score 41;  DB 11;  Length 175;
  Best Local Similarity   58.3%;  Pred. No. 9.1;
  Matches    7;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 ALGILNNQEGWK 12
              :|| | |: |||
Db        110 SLGSLPNENGWK 121


RESULT   9
Q67483
ID   Q67483      PRELIMINARY;      PRT;   905 AA.
AC   Q67483;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   NONSTRUCTURAL PROTEIN 5 (FRAGMENT).
GN   NS5.
OS   Banzi virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC   Flavivirus.
OX   NCBI_TaxID=38837;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SAH 336;
RX   MEDLINE=96005095, PubMed=7561771
RA   Fulop L.D., Barrett A.D., Titball R.W.
RT   "Nucleotide sequence of the NS5 gene of Banzi virus: comparison with
RT   other flaviviruses.";
RL   J. Gen. Virol. 76:2317-2321(1995).
DR   EMBL; L40951, AAA99481.1; -.
DR   InterPro; IPR000208; -.
DR   Pfam; PF00972; Flavi_NS5; 1.
KW   Nonstructural protein.
FT   NON_TER       1
SQ   SEQUENCE   905 AA;  102905 MW;  [illegible]2E08566[illegible] CRC64;

  Query Match             62.1%;  Score 41;  DB 14;  Length 905;
  Best Local Similarity   [illegible]%;  Pred. No. [illegible];
  Matches    7;  Conservative [illegible];  Mismatches [illegible];  Indels [illegible];  Gaps [illegible];

Qy          3 GIENNQ[illegible]GW [illegible]
              |  [illegible]
Db        112 [illegible]TEQ[illegible]W 1[illegible]


RESULT 1[illegible]
Q22[illegible]31
ID   Q22531     PRELIMINARY;     PRT;   [illegible] AA.
AC   Q22531;
DT   01-NOV-1996 (TrEMBLrel. [illegible], Created)
DT   01-OCT-2000 (TrEMBLrel. [illegible], Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   T16G12.[illegible] PROTEIN.
GN   T16G12.[illegible].
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; [illegible]
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Thomas K.;
RL   Submitted (MAR-1994) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; Z30317; CAA829[illegible]
DR   InterPro; IPR00[illegible]30;
DR   InterPro; IPR[illegible]7;
DR   InterPro; IPR00[illegible]30;
DR   Pfam; PF014[illegible]; Peptidase_M1; [illegible]
DR   PRINTS; PR0[illegible]; ALA[illegible]
DR   PROSITE; PS00[illegible]; MI[illegible]; UNKNOWN_[illegible]
DR   PROSITE; PS00[illegible]; ZINC_PROTEASE; UNKNOWN_[illegible]
SQ   SEQUENCE   [illegible] AA;  [illegible] MW;  [illegible] CRC64;

  Query Match             [illegible]%;  Score [illegible];  DB [illegible];  Length [illegible];
  Best Local Similarity   [illegible]%;  Pred. No. [illegible];
  Matches    7;  Conservative [illegible];  Mismatches [illegible];  Indels [illegible];  Gaps [illegible];

Qy          2 [illegible]GIENNQ[illegible]W
              |   ||:  [illegible]
Db       1127 [illegible]GENNRKGW 136


RESULT 11
Q9E[illegible]B
ID   Q9EIV8     PRELIMINARY;     PRT;   [illegible] AA.
AC   Q9EIV8;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   NS5 PROTEIN (FRAGMENT).
GN   NS5.
OS   Edge Hill virus.
OC   Viruses; ssRNA positive-strand viruses, [illegible]
OC   Flavivirus.
OX   NCBI_TaxID=64[illegible];
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=V366;
RA   Macdonald J., Hall R.A., Mackenzie J.[illegible]
RT   "Confirmation of the phylogeny of Edge Hill [illegible]
RT   identification of two Edge Hill [illegible] Australia."
RL   Submitted (JUN-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF275874; AAG17[illegible]
FT   NON_TER       [illegible]     [illegible]
FT   NON_TER     329    329
SQ   SEQUENCE   329 AA;  [illegible] MW;  [illegible] CRC64;

  Query Match             [illegible]%;  Score 1[illegible];  DB [illegible];  Length [illegible];
  Best Local Similarity   [illegible]%;  Pred. No. [illegible];
```

```
  Matches [illegible];  Conservative [illegible];  Mismatches [illegible];  Indels [illegible];  Gaps [illegible];

Qy          [illegible] A[illegible]IENNQ[illegible]WR [illegible]

Db        [illegible] A[illegible]NMEEQ[illegible]WK [illegible]


RESULT [illegible]
Q9EIV[illegible]
ID   [illegible]     PRELIMINARY;     PRT;   [illegible] AA.
AC   [illegible]
DT   [illegible]-MAR-[illegible] (TrEMBLrel. [illegible], Created)
DT   [illegible]-MAR-[illegible] (TrEMBLrel. [illegible], Last sequence update)
DT   [illegible]-MAR-[illegible] (TrEMBLrel. [illegible], Last annotation update)
DE   N[illegible] PROTEIN (FRAGMENT).
GN   N[illegible]
OS   Edge Hill virus.
OC   Viruses; ssRNA [illegible] NA stage; Flaviviridae;
OC   Flavivirus.
OX   NCBI_TaxID=64[illegible];
RN   [illegible]
RP   SEQUENCE FROM N.A.
RC   STRAIN=[illegible];
RA   Macdonald J., Hall R.A., Mackenzie J.[illegible] M.;
RT   "Confirmation of the phylogeny of Edge Hill [illegible] and the
RT   identification of two Edge Hill [illegible] Australia."
RL   Submitted ([illegible]) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF[illegible]; AA[illegible]
FT   NON_TER       [illegible]
FT   NON_TER     [illegible]
SQ   SEQUENCE   [illegible] AA;  [illegible] MW;  [illegible] CRC64;

  Query Match             [illegible]%;  Score 4[illegible];  DB [illegible];  Length [illegible];
  Best Local Similarity   [illegible]%;  Pred. No. [illegible];
  Matches [illegible];  Conservative [illegible];  Mismatches [illegible];  Indels [illegible];  Gaps [illegible];

Qy          [illegible] AIG[illegible]NNQ[illegible]WR [illegible]

Db        [illegible] A[illegible]NM[illegible]EQ[illegible]WK [illegible]


RESULT [illegible]
[illegible]
ID   [illegible]     PRELIMINARY;     PRT;   [illegible] AA.
AC   [illegible]
DT   [illegible]-MAR-[illegible] (TrEMBLrel. [illegible], Created)
DT   01-MAR-[illegible] (TrEMBLrel. [illegible], Last sequence update)
DT   [illegible]-MAR-[illegible] (TrEMBLrel. [illegible], Last annotation update)
DE   PUTATIVE SUGAR TRANSPORTER PERMEASE.
GN   [illegible]
OS   Streptomyces coelicolor.
OC   Bacteria; [illegible]; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; [illegible]; Streptomycetaceae; Streptomyces.
OX   NCBI_TaxID=[illegible];
RN   [illegible]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A3(2);
RA   [illegible]
RL   Submitted ([illegible]) to the EMBL/GenBank/DDBJ databases.
RN   [illegible]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A3(2);
RA   [illegible] M.A.
RL   Submitted ([illegible]) to the EMBL/GenBank/DDBJ databases.
RN   [illegible]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A3(2);
RX   MEDLINE=[illegible]
RA   [illegible] M., [illegible] H.M., [illegible] A., [illegible]
RA   [illegible]
RT   "A[illegible]
RT   [illegible] A3(2)[illegible]"
```

```
RL   Mol. Microbiol. 21:77-96(1996).
DR   EMBL; AL450165; CAC16436.1; -.
KW   Sugar transport.
SQ   SEQUENCE   331 AA;  36866 MW;  4053F3 4B732C68 CRC64;


  Query Match             60.6%;  Score 40;  DB 2;  Length 331;
  Best Local Similarity   63.6%;  Pred. No. 2';
  Matches    7;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 LGILNNQEGWK 12
           ||:|  :||||
Db     285 IGIIYEEGWK 295


RESULT  14
Q31773
ID   Q31773      PRELIMINARY;      PRT;   695 AA.
AC   Q31773;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
     NADH DEHYDROGENASE SUBUNIT (FRAGMENT)
     NDHF.
OS   Atropa belladonna (belladonna).
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I;
OC   Solanales; Solanaceae; Atropa.
OX   NCBI_TaxID=33113;
RN   [1]
RP   SEQUENCE OF 1-436 FROM N.A.
RA   Olmstead R.G., Bremer B., Scott K.M., Palmer J.D.;
RT   "A parsimony analysis of the Asteridae sensu lato based on rbcL
RT   sequences.";
RL   Ann. Mo. Bot. Gard. 80:700-722(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Olmstead R.G., Sweere J.A.;
RT   "Combining data in phylogenetic systematics: an empirical approach
RT   using three molecular data sets in the Solanaceae.";
RL   Syst. Biol. 43:467-481(1994).
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Olmstead R.G.;
RL   Submitted (APR-1994) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U08915; AAA18597.2; -.
DR   InterPro; IPR001516; -.
DR   InterPro; IPR001750; -.
DR   InterPro; IPR002128; -.
     Pfam; PF00361; oxidored_q1; 1.
     Pfam; PF00662; oxidored_q1_N; 1.
DR   Pfam; PF01010; oxidored_q1_C; 1.
KW   Chloroplast.
FT   NON_TER       1      1
FT   NON_TER     695    695
SQ   SEQUENCE   695 AA;  78723 MW;  62F93A25170DC6E7 CRC64;


  Query Match             60.6%;  Score 40;  DB 8;  Length 695;
  Best Local Similarity   63.6%;  Pred. No. 60;
  Matches    7;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 ALGILNNQEGW 11
           ::||  ||||
Db     550 SIGIPFNQEGW 560


RESULT  15
O00851
ID   O00851      PRELIMINARY;      PRT;   1049 AA.
AC   O00851;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   UNCONVENTIONAL MYOSIN IB.
OS   Entamoeba histolytica.
OC   Eukaryota; Entamoebidae; Entamoeba
OX   NCBI_TaxID=5759.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=HM1:IMSS;
RX   MEDLINE=97321558; PubMed=9178268;
RA   Vargas M., Voigt H., Sansonetti P., Guillen N.;
RT   "Molecular characterization of myosin IB from the lower eukaryote
RT   Entamoeba histolytica, a human parasite.";
RL   Mol. Biochem. Parasitol. 86:61-73(1997)
DR   EMBL; U89655; AAC47535.1; -.
DR   HSSP; P08799; 1MND.
DR   InterPro; IPR001452; -.
DR   InterPro; IPR001609; -.
DR   Pfam; PF00018; SH3; 1.
DR   Pfam; PF00063; myosin_head; 1.
DR   PRINTS; PR00193; MYOSINHEAVY.
DR   PRINTS; PR00452; SH3DOMAIN.
DR   ProDom; PD000355; ; 1.
DR   PROSITE; PS50002; SH3; 1.
DR   SMART; SM00326; SH3; 1.
SQ   SEQUENCE   1049 AA;  119551 MW;  213CF358CCAE721B CRC64;


  Query Match             60.6%;  Score 40;  DB 9;  Length 1049;
  Best Local Similarity   77.8%;  Pred. No. 94;
  Matches    7;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 GILNNQEGW 11
           | || |||
Db    1032 GELNGQEGW 1040


Search completed: August 22, 2001, 14:27:14
Job time: 198 sec
```

```
                     GenCore version 4.5
          Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on:          August 22, 2001, 14:29:09 ; Search time 161.96 Seconds
                                            (without alignments)
                                            11.297 Million cell updates/sec

Title:           US-08-659-254-5
Perfect score:   103
Sequence:        1 GSTCVLAGMAIHFGEMPEQ 19

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters:       412676

   mum DB seq length: 0
   mum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       A_Geneseq_0601:*
                1:  /SIDS8/gcgdata/geneseq/geneseqp/AA1980.DAT:*
                2:  /SIDS8/gcgdata/geneseq/geneseqp/AA1981.DAT:*
                3:  /SIDS8/gcgdata/geneseq/geneseqp/AA1982.DAT:*
                4:  /SIDS8/gcgdata/geneseq/geneseqp/AA1983.DAT:*
                5:  /SIDS8/gcgdata/geneseq/geneseqp/AA1984.DAT:*
                6:  /SIDS8/gcgdata/geneseq/geneseqp/AA1985.DAT:*
                7:  /SIDS8/gcgdata/geneseq/geneseqp/AA1986.DAT:*
                8:  /SIDS8/gcgdata/geneseq/geneseqp/AA1987.DAT:*
                9:  /SIDS8/gcgdata/geneseq/geneseqp/AA1988.DAT:*
                10:  /SIDS8/gcgdata/geneseq/geneseqp/AA1989.DAT:*
                11:  /SIDS8/gcgdata/geneseq/geneseqp/AA1990.DAT:*
                12:  /SIDS8/gcgdata/geneseq/geneseqp/AA1991.DAT:*
                13:  /SIDS8/gcgdata/geneseq/geneseqp/AA1992.DAT:*
                14:  /SIDS8/gcgdata/geneseq/geneseqp/AA1993.DAT:*
                15:  /SIDS8/gcgdata/geneseq/geneseqp/AA1994.DAT:*
                16:  /SIDS8/gcgdata/geneseq/geneseqp/AA1995.DAT:*
                17:  /SIDS8/gcgdata/geneseq/geneseqp/AA1996.DAT:*
                18:  /SIDS8/gcgdata/geneseq/geneseqp/AA1997.DAT:*
                19:  /SIDS8/gcgdata/geneseq/geneseqp/AA1998.DAT:*
                20:  /SIDS8/gcgdata/geneseq/geneseqp/AA1999.DAT:*
                21:  /SIDS8/gcgdata/geneseq/geneseqp/AA2000.DAT:*
                22:  /SIDS8/gcgdata/geneseq/geneseqp/AA2001.DAT:*

    Pred. No. is the number of results predicted by chance to have a
    score greater than or equal to the score of the result being printed,
    and is derived by analysis of the total score distribution.

                              SUMMARIES
              %
Result        Query
  No.  Score  Match Length DB  ID                    Description
-----------------------------------------------------------------------------
    1    103  100.0     19  20  AAW74156              Mouse steroidogeni
    2    103  100.0     19  22  AAY728[illegible]     Mouse tryptic pept
    3    103  100.0    284  20  AAW741[illegible]     Mouse steroidogeni
    4    103  100.0    284  22  AAY728[illegible]     Mouse steroidogeni
    5     89   86.4    285  22  AAY728[illegible]     Human steroidogeni
    6     83   80.6    284  17  AAW1323[illegible]    Human steroidogene
    7     83   80.6    285  17  AAW[illegible]323[illegible]  Human steroidogene
    8     83   80.6    285  17  AAW[illegible]323[illegible]  Human steroidogene
    9     83   80.6    285  17  AAW[illegible]081[illegible]  Human steroidogene
   10     83   80.6    285  17  AAW132[illegible]     Human steroidogene
   11     83   80.6    285  17  AAW[illegible]32[illegible]   Human steroidogene
   12     83   80.6    285  17  AAW1[illegible]       Human steroidogene
   13     44   42.7    433  17  AAR984[illegible]1    Oligodendrocyte-my
   14     44   42.7    433  21  AAY[illegible]        Human oligodendroc
   15     43   41.7    275  21  AAB[illegible]        Human mesenchymal
   16     42   40.8    179  19  AAW75[illegible]      Human secreted pro
   17     42   40.8    383  22  AAB75[illegible]      Drosophila glutato
   18     40   38.8    271  21  AAC414[illegible]     Arabidopsis thalia
   19     40   38.8    294  21  AAC41[illegible]      Arabidopsis thalia
   20     40   38.8    387  21  AAC41[illegible]54    Arabidopsis thalia
   21     40   38.8    402   9  AAP[illegible]       Sequence of human
   22     40   38.8    402  11  AAP07[illegible]86    Plasminogen Activa
   23     40   38.8    402  13  AAP23[illegible]      PAI 1 (Arg [illegible]) mu
   24     40   38.8    402  13  AAP23[illegible]      PAI-1 (Arg 351) mu
   25     40   38.8    402  18  AAW[illegible]        Human plasminogen
   26     40   38.8    439  22  AAB76[illegible]      Corynebacterium gl
   27     40   38.8    543  13  AAP28[illegible]      CR1 4 (364-367 NAA
   28     40   38.8    837  20  AAW97[illegible]      Streptococcus pneu
   29     40   38.8   2439  21  AAY58[illegible]      Sorangium cellulos
   30   39.5   38.3    420  20  AAY21[illegible]34    Ligand binding dom
   31   39.5   38.3    441  14  AAR41[illegible]      Steroid hormone re
   32   39.5   38.3    441  17  AAR49[illegible]      Peroxisome prolife
   33   39.5   38.3    441  20  AAY[illegible]        Human PPAR-delta p
   34     39   37.9     62  13  AAR28[illegible]      Human complement t
   35     39   37.9     62  16  AAR72[illegible]      CR1 SCR-[illegible]. Homo s
   36     39   37.9     62  19  AAW66[illegible]      Short consensus re
   37     39   37.9     62  20  AAY55[illegible]      Human CR1 protein
   38     39   37.9     62  22  AAB29[illegible]      Human CR1 compleme
   39     39   37.9    109  21  AAG54[illegible]      Arabidopsis thalia
   40     39   37.9    145  21  AAY944[illegible]     Protein monomer of
   41     39   37.9    438  20  AAY55756              Human CR1 protein
   42     39   [illegible]  [illegible]  22  AA[illegible]  Human PRO15 protei
   43     39   [illegible]  156  20  AAY55[illegible]  Human CR1 protein
   44     39   [illegible]  [illegible]  21  AA[illegible]  Streptomyces globi
   45     39   37.9    616  20  AAY[illegible]        Drosophila melanog

                              ALIGNMENTS

RESULT      1
AAW74156
ID   AAW74156 standard; Protein; 19 AA.
XX
AC   AAW74156;
XX
DT   05-MAY-1999  (first entry)
XX
DE   Mouse steroidogenic acute regulatory protein tryptic peptide.
XX
KW   StAR; steroidogenic acute regulatory protein; McCune-Albright syndrome;
KW   steroid hormone dependent disorder; adrenal hyperplasia congenita; LCAH;
KW   lipoid congenital adrenal hyperplasia; infertility; sexual maturation;
KW   androgen-responsive tumours; hypogonadism; therapy; progesterone level.
XX
OS   Mus sp.
XX
PN   US5872230 A.
XX
PD   16-FEB-1999.
XX
PF   04-NOV-1995;    95US-0538960.
XX
PR   04-NOV-1995;    95US-0538960.
XX
PA   (UYTE ) UNIV TEXAS TECH HEALTH SCI CENT
XX
PI   Clark BJ,  Stocco DM;
XX
DR   WPI; 1999-166726/14.
XX
PT   New nucleic acid encoding steroidogenic acute regulatory protein -
PT   useful for treatment of steroid hormone-dependent disorders such as
PT   lipoid congenital adrenal hyperplasia, infertility and
PT   androgen responsive tumours
XX
PS   Example 2; Column 29; 27pp; English.
XX
CC   This sequence is a tryptic peptide fragment [illegible]
CC   regulatory protein (StAR) of the invention. The StAR [illegible] sequence and
CC   its products or derivatives may be used as probes [illegible]
CC   hybridisation studies and in the treatment of steroid hormone dependent
CC   disorders such as lipoid congenital adrenal hyperplasia (LCAH),
CC   infertility, sexual maturation, androgen responsive tumours, [illegible]
CC   puberty, McCune Albright syndrome, adrenal hypoplasia [illegible]
CC   hypogonadism. Maternal levels of StAR may [illegible]
CC   progesterone levels and so a risk of spontaneous [illegible]
XX
SQ   Sequence 19 AA;

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches 19; Conservative [illegible] Mismatches [illegible]

Qy          1 GSDPVLAGMATHEGEMPEQ 19
                  |||        ||
Db          1 qstpvlaqmatsqeeeq 19


RESULT 2
AAY72834
ID   AAY72834 standard; peptide; 19 AA.
XX
AC   AAY72834;
XX
DT   31 MAY 2001  (first entry)
XX
DE   Mouse tryptic peptide #45 for analysis of [illegible] StAR protein.
XX
KW   Mouse; steroidogenic acute regulatory protein; StAR;
KW   gene therapy; steroid hormone dependent disorder; [illegible]
KW   lipoid congenital adrenal hyperplasia; sexual maturation;
KW   infertility; androgen responsive tumour; [illegible]
KW   hypogonadotropic hypogonadism; McCune Albright [illegible]
KW   adrenal hypoplasia congenita; [illegible]
XX
OS   Mus sp.
XX
PN   US6194[illegible] B1.
XX
PD   27 FEB 2001.
XX
PF   07 JUN 1996; [illegible]
XX
PR   04 NOV 1995; [illegible]
XX
PA   (UYTE ) UNIV TEXAS [illegible] HEALTH [illegible]
XX
PI   Stocco DM, Clark BJ;
XX
DR   WPI; 2001-217906/22.
XX
PT   Novel nucleic acid encoding steroidogenic acute regulatory protein
PT   useful for treating steroid hormone dependent disorders
XX
PS   Example 2; Column 28; 31pp; English.
XX
CC   The invention relates to steroidogenic acute regulatory protein
CC   (StAR) and nucleic acid molecules encoding these StAR proteins,
CC   are used for regulating steroidogenesis. [illegible]
CC   proteins are involved in regulating [illegible]
CC   the mitochondria of a cell. The [illegible]
CC   invention are useful in gene therapy for the treatment of [illegible]
CC   hormone dependent disorders, particularly lipoid congenital
CC   adrenal hyperplasia (LCAH), infertility, sexual maturation,
CC   androgen responsive tumours, precocious puberty, McCune Albright
CC   syndrome, adrenal hypoplasia congenita or hypogonadotropic
CC   hypogonadism. The [illegible] useful in the preparation
CC   of the steroidogenic acute regulatory protein.
CC   The present sequence is tryptic peptide #45. This peptide [illegible]
CC   used in the [illegible] and sequence analysis of cDNA encoding the
CC   [illegible] form of StAR protein. This peptide [illegible] used in the
CC   analysis of the StAR protein sequence. Peptide #45 is useful
CC   for designing oligonucleotides for screening and identifying
CC   antigenic determinants of StAR protein.
XX
SQ   Sequence 19 AA;

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches [illegible] Mismatches [illegible] Indels [illegible] Gaps [illegible]

Qy          [illegible] VLAGMAIHEGEMPEQ [illegible]
                  [illegible]
Db          [illegible] lamatsqerpeq [illegible]


RESULT [illegible]
AAW[illegible]
ID   AAW[illegible] standard; peptide; [illegible] AA.
XX
AC   AAW[illegible]
XX
DT   [illegible] MAY [illegible] (first entry)
XX
DE   Mouse [illegible] acute regulatory protein [illegible]
XX
KW   StAR; steroidogenic acute regulatory protein; McCune Albright syndrome;
KW   steroid hormone dependent disorder; adrenal hypoplasia congenita; LCAH;
KW   lipoid congenital adrenal hyperplasia; infertility; sexual maturation;
KW   androgen responsive tumours; hypogonadism; therapy; progesterone levels.
XX
OS   Mus sp.
XX
PN   [illegible] A.
XX
PD   [illegible] FEB [illegible]
XX
PF   [illegible] NOV [illegible]
XX
PR   [illegible] NOV [illegible]
XX
PA   ([illegible]) UNIV TEXAS [illegible] HEALTH [illegible]
XX
PI   Clark BJ, [illegible] DM;
XX
DR   WPI; [illegible]
DR   N-PSDB; AAX[illegible], AAX[illegible]
XX
PT   New nucleic acid encoding steroidogenic acute regulatory protein
PT   useful for treating [illegible] hormone dependent disorders such as
PT   lipoid congenital adrenal hyperplasia, infertility and
PT   androgen responsive tumours
XX
PS   Claim [illegible]; [illegible]pp; English.
XX
CC   This sequence is the steroidogenic acute regulatory protein (StAR)
CC   of the invention. The StAR [illegible] sequence and its products or
CC   derivatives may be used as probes and primers for PCR hybridisation
CC   [illegible] and in the treatment of steroid hormone dependent disorders such
CC   as lipoid congenital adrenal hyperplasia (LCAH), infertility, sexual
CC   maturation, androgen responsive tumours, precocious puberty,
CC   McCune Albright syndrome, adrenal hypoplasia congenita and hypogonadism.
CC   Maternal levels of StAR may [illegible] progesterone levels
CC   and so a risk of [illegible]
XX
SQ   Sequence [illegible] AA;
```

```
  Query Match             100.0%;  Score [illegible];  DB 20;  Length [illegible];
  Best Local Similarity   100.0%;  Pred. No. 1.2e-08;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 GSTCVLAGMATHFGEMPEQ 19
            |||||||||||||||||||
Db      193 gstcvlagmathfgempeq 211


RESULT    4
AAY72831
ID   AAY72831 standard; Protein; 284 AA.
XX
AC   AAY72831;
XX
DT   31-MAY-2001  (first entry)
XX
DE   Mouse steroidogenic acute regulatory (StAR) protein.
XX
KW   Mouse; steroidogenic acute regulatory protein; StAR;
KW   gene therapy; steroid hormone-dependent disorder; LCAH;
     lipoid congenital adrenal hyperplasia; sexual maturation;
     infertility; androgen-responsive tumour; precocious puberty;
KW   hypogonadotropic hypogonadism; McCune-Albright syndrome;
KW   adrenal-hypoplasia congenita; cytostatic.
XX
OS   Mus sp.
XX
FH   Key              Location/Qualifiers
FT   Peptide          1..47
FT                    /label= Signal_peptide
FT   Cleavage-site    39..40
FT                    /note= "Site for mitochondrial proteases,
FT                    matrix-processing protease and mitochondrial
FT                    intermediate processing peptide"
FT   Protein          48..284
FT                    /label= Mature_steroidogenic_acute_regulatory_protein
FT   Cleavage-site    49..50
FT                    /note= "Site for mitochondrial proteases,
FT                    matrix-processing protease and mitochondrial
FT                    intermediate processing peptide"
FT   Modified-site    59..60
FT                    /note= "PKA phosphorylation site"
FT   Modified-site    68..69
FT                    /note= "CK phosphorylation site"
FT   Region           85..96
FT                    /label= Tryptic_peptide #25
FT   Region           88..98
FT                    /note= "This sequence is used to raise antibodies
                      against StAR protein"
     Region           193..211
                      /label= Tryptic_peptide #45
FT   Modified-site    195
FT                    /note= "PKA phosphorylation site"
FT   Region           217..230
FT                    /label= Tryptic_peptide #23
FT   Modified site    233..234
FT                    /note= "P34CDC2 kinase (C2) phosphorylation site"
FT   Modified-site    238
FT                    /note= "PKC phosphorylation site"
XX
PN   US6194555-B1.
XX
PD   27-FEB-2001.
XX
PF   07-JUN-1996;   96US-0659254.
XX
PR   04-NOV-1995;   95US-0538960.
XX
PA   (UYTE-) UNIV TEXAS TECH HEALTH SCI CENT.
XX
PI   Stocco DM,  Clark BJ;
```

```
XX
DR   WPI; 2001-217[illegible]/22.
DR   N-PSDB; AAD02901, AAD02902.
XX
PT   Novel nucleic acid encoding steroidogenic acute regulatory protein,
PT   useful for treating steroid hormone-dependent disorders
XX
PS   Example 2; Fig 2; 41pp; English.
XX
CC   The invention relates to steroidogenic acute regulatory protein
CC   (StAR) and nucleic acid molecules encoding them. StAR proteins
CC   are used for regulating steroidogenesis. More particularly these
CC   proteins are involved in regulating cholesterol transport into
CC   the mitochondria of a cell. The nucleic acid molecules of the
CC   invention are useful in gene therapy for the treatment of steroid
CC   hormone-dependent disorders, particularly, lipoid congenital
CC   adrenal hyperplasia (LCAH), infertility, sexual maturation,
CC   androgen-responsive tumours, precocious puberty, McCune-Albright
CC   syndrome, adrenal-hypoplasia congenita or hypogonadotropic
CC   hypogonadism. They are also useful for recombinant preparation
CC   of the steroidogenic acute regulatory protein.
CC   The present sequence is mouse steroidogenic acute regulatory
CC   protein (StAR).
XX
SQ   Sequence   284 AA;


  Query Match             100.0%;  Score [illegible];  DB 22;  Length 284;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-08;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 GSTCVLAGMATHFGEMPEQ 19
            |||||||||||||||||||
Db      193 gstcvlagmathfgempeq 211


RESULT    5
AAY72838
ID   AAY72838 standard; Protein; 285 AA.
XX
AC   AAY72838;
XX
DT   31-MAY-2001  (first entry)
XX
DE   Human steroidogenic acute regulatory (StAR) protein.
XX
KW   Human; steroidogenic acute regulatory protein; StAR;
KW   gene therapy; steroid hormone-dependent disorder; LCAH;
KW   lipoid congenital adrenal hyperplasia; sexual maturation;
KW   infertility; androgen-responsive tumour; precocious puberty;
KW   hypogonadotropic hypogonadism; McCune-Albright syndrome;
KW   adrenal hypoplasia congenita; cytostatic.
XX
OS   Homo sapiens.
XX
FH   Key              Location/Qualifiers
FT   Misc-difference 156
FT                    /note= "Encoded by GAC; this feature corresponds
FT                    to the translation exception for the sequence shown
FT                    in (AAD02909)"
FT   Modified site    59..60
FT                    /note= "PKA phosphorylation site"
FT   Modified-site    68..69
FT                    /note= "CK phosphorylation site"
FT   Modified-site    195
FT                    /note= "PKA phosphorylation site"
FT   Modified-site    233..234
FT                    /note= "P34CDC2 kinase (C2) phosphorylation site"
FT   Modified-site    238
FT                    /note= "PKC phosphorylation site"
FT   Modified-site    277
FT                    /note= "PKA phosphorylation site"
XX
```

```
PN   [illegible] B1
XX
PD   2[illegible] FEB 2001.
XX
PF   07 JUN 1999; [illegible]
XX
PR   04 NOV 199[illegible]; [illegible] US [illegible]
XX
PA   (UYTE ) UNIV TEXAS [illegible] HEALTH [illegible]
XX
PI   Stocco DM, [illegible] B;
XX
DR   WPI; 2001-21[illegible]/22.
DR   N-PSDB; AA[illegible]; AA[illegible].
XX
PT   Novel nucleic acid [illegible]
PT   useful for treating steroid hormone dependent disorder
XX
PS   Example 14; [illegible] 4[illegible]pp; English.
XX
CC   The invention relates to steroid[illegible]
CC   (StAR) and nucleic acid molecules encoding them. [illegible]
CC   are used for [illegible] steroid[illegible] M[illegible]
CC   proteins are involved in regulating cholesterol [illegible]
CC   the mitochondria of [illegible] The [illegible]
CC   invention are useful in gene therapy for the treatment of [illegible]
CC   hormone dependent disorders, particularly [illegible]
CC   adrenal hyperplasia (CAH), infertility, [illegible]
CC   androgen responsive tumours, prostate [illegible]
CC   syndromes, adrenal hyp[illegible]
CC   hypogonadism. They are also useful for [illegible]
CC   of the steroid[illegible] acute regulatory protein.
CC   The present sequence is human steroidogenic acute regulatory
CC   protein (StAR)
XX
SQ   Sequence [illegible] AA;

  Query Match             [illegible] 14[illegible]
  Best Local Similarity   [illegible]; Pred. No. [illegible]
  Matches [illegible]; Conservative [illegible]; Mismatches [illegible]

Qy     [illegible] VLAGMA[illegible] SEM[illegible]
          [illegible]
Db    194 [illegible]


RESULT 6
AAW13238
ID   AAW13238 standard; Protein; [illegible] AA.
XX
AC   AAW13238;
XX
DT   21 MAY 1997 [illegible]
XX
DE   Human steroidogenesis acute regulatory [illegible]
XX
KW   Human; steroidogenesis [illegible] acute regulatory protein; [illegible]
KW   mutation; detection; prenatal; genetic defect; [illegible]
KW   lipoid adrenal hyperplasia; treatment; prevention; gene
KW   replacement therapy; [illegible]
XX
OS   Homo sapiens.
XX
PN   WO9629338-A1.
XX
PD   26-SEP-1996.
XX
PF   22-MAR-1996; [illegible]
XX
PR   23-MAR-1995; [illegible]
XX
PA   (REGC ) UNIV CALIFORNIA.
```

```
PA   ([illegible] ) UNIV PENNSYLVANIA.
XX
PI   [illegible] Miller WL, Strauss JF;
XX
DR   WPI; [illegible]
XX
PT   [illegible] regulatory protein gene [illegible] used
PT   for detection of mutation(s) of this gene that cause congenital
PT   lipoid adrenal hyperplasia
XX
PS   [illegible]
XX
CC   The present sequence is the human steroid[illegible] acute regulatory
CC   protein (hStAR) [illegible] lipoid adrenal hyperplasia (CAH)
CC   [illegible] Ara[illegible] has
CC   been [illegible] the hStAR [illegible] mutation is
CC   useful in the detection (e.g. prenatal) of CAH associated
CC   [illegible] CAH can be
CC   treated by protein or gene replacement therapy which can also be
CC   [illegible]
XX
SQ   Sequence [illegible] AA;

  Query Match             [illegible] Length [illegible];
  Best Local Similarity   [illegible]; Pred. No. [illegible]
  Matches [illegible]; Mismatches [illegible]; Gaps [illegible]

Qy     [illegible] A[illegible]MA[illegible] MEL[illegible]
          [illegible]
Db     [illegible]


RESULT [illegible]
AAW[illegible]
ID   AAW[illegible] standard; Protein; [illegible] AA.
XX
AC   AAW[illegible]
XX
DT   [illegible] MAY [illegible]
XX
DE   Human steroid[illegible] regulatory protein mutant Ala2[illegible]Val.
XX
KW   Human; steroidogenesis [illegible] acute regulatory protein; hStAR; [illegible]
KW   [illegible] prenatal; genetic defect; [illegible] protein;
KW   lipoid adrenal hyperplasia; treatment; prevention; gene
KW   replacement therapy; [illegible]
XX
OS   [illegible]
XX
FH   Key             Location/Qualifiers
FT   M[illegible]
FT                   /note= "wild type Ala [illegible] mutated with Val"
XX
PN   WO[illegible]
XX
PD   [illegible]
XX
PF   [illegible] MAR [illegible] W[illegible]
XX
PR   [illegible] MAR [illegible]
XX
PA   ([illegible] ) UNIV CALIFORNIA.
PA   ([illegible] ) UNIV PENNSYLVANIA.
XX
PI   [illegible] M[illegible] W[illegible] [illegible]
XX
DR   WPI; [illegible]/43.
DR   N-PSDB; AA[illegible]
XX
PT   [illegible] regulatory protein gene [illegible] used
PT   [illegible] cause congenital
PT   lipoid adrenal hyperplasia
```

```
XX
PS   Claim 13; Page -; 89pp; English.
XX
CC   The present sequence is the human steroidogenesis acute regulatory
CC   protein (hStAR), congenital lipoid adrenal hyperplasia (CAH)
CC   associated mutant, Ala218Val. Analysing the hStAR gene for such a
CC   mutation is useful in the detection (e.g. prenatally) of CAH
CC   associated genetic defects, or its transmission to children. CAH
CC   can be treated by protein or gene replacement therapy, which can
CC   also be used to prevent or treat hypercholesterolaemia.
XX
SQ   Sequence   285 AA;


  Query Match             80.6%;  Score 83;  DB 17;  Length 285;
  Best Local Similarity   84.2%;  Pred. No. 1.8e-05;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy        1 GSTCVLAGMATHFGEMPEQ 19
              ||||||||  | || ||||
Db      194 qstcvlaqmdtdtqnmpeq 212

      LT   8
AAW13237
ID   AAW13237 standard; Protein; 285 AA.
XX
AC   AAW13237;
XX
DT   21-MAY-1997  (first entry)
XX
DE   Human steroidogenesis acute regulatory protein mutant Leu275Pro.
XX
KW   Human; steroidogenesis; acute regulatory protein; hStAR; analysis;
KW   mutation; detection; prenatal; genetic defect; congenital; protein;
KW   lipoid adrenal hyperplasia; treatment; prevention; gene;
KW   replacement therapy; hypercholesterolaemia; mutant.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 275
FT                   /note= "wild type Leu substituted with Pro"
XX
PN   WO9629338-A1.
XX
PD   26-SEP-1996.
XX
PF   22-MAR-1996;   96WO-US03896.
XX
     23-MAR-1995;   95US-0410540.
XX
     (REGC ) UNIV CALIFORNIA.
PA   (UYPE-) UNIV PENNSYLVANIA.
XX
PI   Lin D,  Miller WL,  Strauss JF;
XX
DR   WPI; 1996-443130/44.
DR   N-PSDB; AAT62022.
XX
PT   Isolated human steroidogenesis acute regulatory protein gene - used
PT   for detection of mutation(s) of this gene that cause congenital
PT   lipoid adrenal hyperplasia
XX
PS   Claim 13; Page -; 89pp; English.
XX
CC   The present sequence is the human steroidogenesis acute regulatory
CC   protein (hStAR), congenital lipoid adrenal hyperplasia (CAH)
CC   associated mutant, Leu275Pro. Analysing the hStAR gene for such a
CC   mutation is useful in the detection (e.g. prenatally) of CAH
CC   associated genetic defects, or its transmission to children. CAH
CC   can be treated by protein or gene replacement therapy, which can
CC   also be used to prevent or treat hypercholesterolaemia.
```

```
XX
SQ   Sequence   285 AA;


  Query Match             80.6%;  Score 83;  DB 17;  Length 285;
  Best Local Similarity   84.2%;  Pred. No. 1.8e-05;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy        1 GSTCVLAGMATHFGEMPEQ 19
              ||||||||  | || ||||
Db      194 qstcvlaqmdtdtqnmpeq 212


RESULT   9
AAW00814
ID   AAW00814 standard; Protein; 285 AA
XX
AC   AAW00814;
XX
DT   21 MAY-1997  (first entry)
XX
DE   Human steroidogenesis acute regulatory protein.
XX
KW   Human; steroidogenesis; acute regulatory protein; hStAR; analysis;
KW   mutation; detection; prenatal; genetic defect; congenital; protein;
KW   lipoid adrenal hyperplasia; treatment; prevention; gene;
KW   replacement therapy; hypercholesterolaemia.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Region          5
FT                   /note= "potential site for protein kinase C
FT                          mediated phosphorylation"
FT   Region          13
FT                   /note= "potential site for protein kinase C
FT                          mediated phosphorylation"
FT   Region          56..57
FT                   /note= "potential site for protein kinase A
FT                          mediated phosphorylation"
FT   Region          100
FT                   /note= "potential site for protein kinase A
FT                          mediated phosphorylation"
FT   Region          110
FT                   /note= "potential site for protein kinase A
FT                          mediated phosphorylation"
FT   Region          186
FT                   /note= "potential site for protein kinase C
FT                          mediated phosphorylation"
FT   Region          195..196
FT                   /note= "potential site for protein kinase A
FT                          mediated phosphorylation"
FT   Region          277
FT                   /note= "potential site for protein kinase A
FT                          mediated phosphorylation"
XX
PN   WO9629338 A1.
XX
PD   26-SEP-1996.
XX
PF   22-MAR-1996;   96WO-US03896.
XX
PR   23-MAR-1995;   95US-0410540.
XX
PA   (REGC ) UNIV CALIFORNIA.
PA   (UYPE ) UNIV PENNSYLVANIA.
XX
PI   Lin D,  Miller WL,  Strauss JF
XX
DR   WPI; 1996-443130/44.
DR   N PSDB; AAT33650.
XX
PT   Isolated human steroidogenesis acute regulatory protein gene - used
```

```
PI   for detection of mutation(s) of this gene that [illegible]
PI   lipoid adrenal hyperplasia
XX
PS   Claim 1; Page 63; 80pp; English.
XX
CC   The present sequence is the human [illegible]
CC   regulatory protein (hStAR). The hStAR [illegible]
CC   mutations to detect [illegible] (prenatally) [illegible]
CC   with congenital lipoid adrenal hyperplasia (CAH) [illegible]
CC   transmission to children. CAH can be treated [illegible]
CC   replacement therapy, which can also be used [illegible]
CC   hypercholesterolemia
CC   A human adrenal cortex cDNA library was [illegible] StAR
CC   probe to isolate a [illegible] kb insert [illegible]
CC   residue protein. When it was [illegible]
CC   COS-1 cells cotransfected with p[illegible]
CC   level of pregnenolone synthesis from [illegible]
CC   20 alpha hydroxycholesterol.
XX
SQ   Sequence 285 AA;

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches 16; Conservative [illegible] Mismatches [illegible]

Qy       [illegible]VLAGMA[illegible]SEM[illegible]
             [illegible]
Db     [illegible]vlagmA[illegible]

RESULT 10
AAW13233
ID   AAW13233 standard; Protein; [illegible] AA.
XX
AC   AAW13233;
XX
DT   21 MAY 1997 (first entry)
XX
DE   Human steroidogenesis acute regulatory protein [illegible]
XX
KW   Human; steroidogenesis; acute regulatory protein; [illegible]
KW   mutation; detection; prenatal; genetic defect; [illegible]
KW   lipoid adrenal hyperplasia; treatment; [illegible]
KW   replacement therapy; hypercholesterolaemia; [illegible]
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference [illegible]
FT                   /note= "wild type [illegible]"
XX
PN   WO9629[illegible]-A1.
XX
PD   26-SEP-1996.
XX
PF   22-MAR-1996;   [illegible]
XX
PR   23-MAR-1995;   [illegible]
XX
PA   (REGC ) UNIV CALIFORNIA.
PA   (UYPE ) UNIV PENNSYLVANIA.
XX
PI   Lin D,  Miller WL,  [illegible]
XX
DR   WPI; 1996-44[illegible]
DR   N-PSDB; AAT62018.
XX
PT   Isolated human steroidogenesis acute regulatory [illegible]
PT   for detection of mutation(s) of this gene that [illegible]
PT   lipoid adrenal hyperplasia
XX
PS   Claim 1[illegible]; Page [illegible]; [illegible]pp; English.
```

```
XX
CC   [illegible] regulatory
CC   [illegible] (hStAR), congenital lipoid adrenal hyperplasia (CAH)
CC   [illegible] Analysing the hStAR gene for [illegible]
CC   [illegible] the detection (e.g. prenatally) of CAH
CC   associated genetic defects or its transmission to children. CAH
CC   can be treated [illegible] replacement therapy, which can
CC   [illegible] hypercholesterolaemia.
XX
SQ   [illegible] AA;

  Query Match [illegible] Length [illegible];
  Best Local Similarity [illegible]; Pred. No. [illegible]
  Matches [illegible] Mismatches [illegible] Indels [illegible] Gaps [illegible]

Qy       [illegible]VLAGMA[illegible]
             [illegible]
Db     [illegible]

RESULT 11
AAW[illegible]
ID   AAW[illegible] standard; Protein; [illegible] AA.
XX
AC   AAW[illegible];
XX
DT   [illegible] MAY [illegible]
XX
DE   Human steroidogenesis acute regulatory protein mutant [illegible].
XX
KW   Human; steroidogenesis; acute regulatory protein; hStAR; analysis;
KW   mutation; detection; prenatal; genetic defect; congenital; [illegible]
KW   lipoid adrenal hyperplasia; treatment; [illegible]
KW   replacement therapy; hypercholesterolaemia; mutant.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference [illegible]
FT                   /note= "wild type [illegible] with [illegible]"
XX
PN   WO9629[illegible]-A1.
XX
PD   26-SEP-1996.
XX
PF   22-MAR-1996;   [illegible]
XX
PR   23-MAR-1995;   [illegible]
XX
PA   (REGC ) UNIV CALIFORNIA.
PA   (UYPE ) UNIV PENNSYLVANIA.
XX
PI   Lin D,  Miller WL,  [illegible]
XX
DR   WPI; 1996-[illegible]
DR   N-PSDB; AA[illegible]
XX
PT   Isolated human steroidogenesis acute regulatory protein gene - used
PT   for detection of mutation(s) of this gene that cause congenital
PT   lipoid adrenal hyperplasia
XX
PS   [illegible]
XX
CC   The present sequence is the human steroidogenesis acute regulatory
CC   protein (hStAR), congenital lipoid adrenal hyperplasia (CAH)
CC   associated mutant. Analysing the hStAR gene for [illegible]
CC   mutations is useful in the detection (e.g. prenatally) of CAH
CC   associated genetic defects or its transmission to children. CAH
CC   can be treated [illegible] replacement therapy, which can
CC   [illegible] hypercholesterolaemia.
XX
SQ   [illegible] AA;
```

```
  Query Match             80.6%;  Score 83;  DB 17;  Length 285;
  Best Local Similarity   84.2%;  Pred. No. 1.8e-05;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    :

Qy        1 GSTCVLAGMATHFGEMPEQ 19
            ||||||||| |   | ||||
Db      194 qstcvlaqmdtdfqnmpeq 212


RESULT  12
AAW13235
ID   AAW13235 standard; Protein; 285 AA.
XX
AC   AAW13235;
XX
DT   21-MAY-1997  (first entry)
XX
DE   Human steroidogenesis acute regulatory protein mutant Glu169Lys.
XX
KW   Human; steroidogenesis; acute regulatory protein; hStAR; analysis;
[illegible] mutation; detection; prenatal; genetic defect; congenital; protein;
[illegible] lipoid adrenal hyperplasia; treatment; prevention; gene;
KW   replacement therapy; hypercholesterolaemia; mutant.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 169
FT                   /note= "wild type Glu substituted with Lys"
XX
PN   WO9629338-A1.
XX
PD   26-SEP-1996.
XX
PF   22-MAR-1996;   96WO-US03896.
XX
PR   23-MAR-1995;   95US-0410540.
XX
PA   (REGC ) UNIV CALIFORNIA.
PA   (UYPE-) UNIV PENNSYLVANIA.
XX
PI   Lin D,  Miller WL,  Strauss JF;
XX
DR   WPI; 1996-443130/44.
DR   N-PSDB; AAT62020.
XX
PT   Isolated human steroidogenesis acute regulatory protein gene - used
PT   for detection of mutation(s) of this gene that cause congenital
[illegible] lipoid adrenal hyperplasia
XX
PS   Claim 13; Page  ; 89pp; English.
XX
CC   The present sequence is the human steroidogenesis acute regulatory
CC   protein (hStAR), congenital lipoid adrenal hyperplasia (CAH)
CC   associated mutant, Glu169Lys. Analysing the hStAR gene for such a
CC   mutation is useful in the detection (e.g. prenatally) of CAH
CC   associated genetic defects, or its transmission to children. CAH
CC   can be treated by protein or gene replacement therapy, which can
CC   also be used to prevent or treat hypercholesterolaemia.
XX
SQ   Sequence   285 AA;


  Query Match             80.6%;  Score 83;  DB 17;  Length 285;
  Best Local Similarity   84.2%;  Pred. No. 1.8e-05;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy        1 GSTCVLAGMATHFGEMPEQ 19
            ||||||||| | || ||||
Db      194 qstcvlaqmdtdfqnmpeq 212
```

```
RESULT  13
AAR98454
ID   AAR98454 standard; Protein; 433 AA
XX
AC   AAR98454;
XX
DT   15-SEP-1996  (first entry)
XX
DE   Oligodendrocyte-myelin glycoprotein
XX
KW   Oligodendrocyte myelin glycoprotein; OMgp; CR-LR protein;
KW   central nervous system; CNS; myelination; neurofibromatosis;
KW   diagnosis.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..17
FT                   /label= Sig_peptide
FT   Misc-difference 14
FT                   /label= Gly, Asp
FT                   /note= "ambiguity in the nucleic acid sequence"
FT   Domain          18..49
FT                   /label= N terminal_EGF_motif
FT   Domain          50..221
FT                   /label= Leucine rich repeat_domain
FT   Region          50..72
FT                   /note= "leucine-rich repeat"
FT   Region          73..93
FT                   /note= "leucine-rich repeat"
FT   Region          94..117
FT                   /note= "leucine-rich repeat"
FT   Region          118..140
FT                   /note= "leucine-rich repeat"
FT   Region          141..161
FT                   /note= "leucine-rich repeat"
FT   Region          162..185
FT                   /note= "leucine-rich repeat"
FT   Region          186..209
FT                   /note= "leucine-rich repeat"
FT   Region          210..221
FT                   /note= "leucine-rich repeat"
FT   Domain          222..409
FT                   /label= Serine threonine repeat_domain
FT   Region          222..263
FT                   /note= "serine-threonine repeat"
FT   Region          264..285
FT                   /note= "serine-threonine repeat"
FT   Region          286..328
FT                   /note= "serine-threonine repeat"
FT   Region          329..370
FT                   /note= "serine-threonine repeat"
FT   Region          371..409
FT                   /note= "serine-threonine repeat"
FT   Domain          410..418
FT                   /label= C-terminal domain
FT   Peptide         419..433
FT                   /label= Sig_peptide
FT                   /note= "putative C-terminal signal peptide"
XX
PN   US5532351-A.
XX
PD   02-JUL-1996.
XX
PF   12-JUL-1990;   90US-0551267.
XX
PR   12-JUL-1990;   90US-0551267.
PR   27-JAN-1994;   94US-0183588.
XX
PA   (ARCH-) ARCH DEV CORP.
XX
PI   Stefansson K;
```

```
XX
DR   WPI; [illegible]
DR   N-PSDB; AA[illegible]
XX
PT   Nucleic acid [illegible]
PT   for diagnosis of neuroblastoma [illegible]
XX
PS   Claim [illegible]
XX
CC   Oligodendrocyte [illegible]
CC   CRLF (cysteine [illegible]
CC   myelination in the CNS. Its [illegible]
CC   acid sequence was derived from a cDNA [illegible]
CC   neonatal spinal cord cDNA library [illegible]
CC   chromosome 17 and maps within [illegible]
CC   breakpoint that [illegible]
CC   The protein may [illegible]
CC   detection of CRLF proteins
XX
SQ   Sequence [illegible] AA;

  Query Match [illegible]
  Best Local Similarity [illegible] Pred. No. [illegible]
  Matches [illegible]

Qy   [illegible] VIAGMA[illegible]
                |        :|
Db   3[illegible] [illegible]

RESULT 14
AAY57091
ID   AAY57091 standard; protein; [illegible] AA.
XX
AC   AAY57091;
XX
DT   28-FEB-2000  (first entry)
XX
DE   Human oligodendrocyte myelin glycoprotein fragment
XX
KW   Neglected target tissue antigen; NTTA; [illegible]
KW   immunotherapeutic agent; insulin dependent diabetes mellitus;
KW   multiple sclerosis; autoimmune thyroiditis; [illegible]
KW   osteoarthritis; inflammatory response.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference [illegible]
FT                   /label= [illegible]
XX
PN   WO99[illegible]-A1.
XX
PD   11-NOV-1999.
XX
PF   07-MAY-1999; [illegible]
XX
PR   07-MAY-1998; [illegible]
XX
PA   (REGC ) UNIV CALIFORNIA.
XX
PI   Kaufman DL, [illegible]
XX
DR   WPI; 2000-[illegible]
XX
PT   Administration of neglected target tissue [illegible]
PT   responses -
XX
PS   Disclosure; Page [illegible]
XX
CC   Amino acid sequences AAY57063-Y57[illegible]
CC   tissue antigens (NTTAs). NTTAs [illegible]
CC   [illegible] These peptides and proteins are used in
CC   the method [illegible] administering an NTTA as an
CC   antigen based immunotherapeutic agent [illegible] afflicted with an
CC   autoimmune [illegible] associated with an autoimmune disease. The
CC   immunotherapeutic agent is used to treat autoimmune diseases such as
CC   insulin dependent diabetes mellitus, multiple sclerosis, autoimmune
CC   thyroiditis, [illegible]
CC   inflammatory immune response. The NTTA induces regulatory [illegible] by
CC   elicitation of regulatory T cells among T cells recognizing the NTTA but
CC   not participating in the disease response. The NTTA are capable of
CC   recruiting a substantial population of uncommitted T cells which can
CC   be primed [illegible] towards regulatory responses to provide effective
CC   treatment. The NTTA are effective [illegible]
CC   responses even when target determinants such as agents promoting
CC   tolerance [illegible] have failed [illegible]
CC   response. NTTAs [illegible] anticipated [illegible] be safer
CC   than use of target determinants.
XX
SQ   Sequence [illegible] AA;

  Query Match [illegible]
  Best Local Similarity [illegible] Pred. No. [illegible]
  Matches [illegible]

Qy   [illegible] VIAGMA[illegible]

Db   [illegible]

RESULT 15
AAB[illegible]
ID   AAB[illegible] standard; protein; [illegible] AA.
XX
AC   AAB[illegible]
XX
DT   [illegible]-FEB-[illegible]  (first entry)
XX
DE   Human mesenchymal stem cell polypeptide
XX
KW   Mesenchymal stem cell; cDNA; detection; [illegible] marker; [illegible]
KW   genetic marker; [illegible] analysis; [illegible] chromosome marker;
KW   bone tissue; therapy; genetic fingerprinting; peptide; immunogen;
KW   [illegible]
XX
OS   Homo sapiens.
XX
PN   WO[illegible]-A2.
XX
PD   [illegible]
XX
PF   [illegible]-MAR-[illegible]
XX
PR   [illegible]-APR-[illegible]
PR   [illegible]-AUG-[illegible]
XX
PA   ([illegible]) [illegible] THERAPEUTICS INC.
XX
PI   [illegible]
XX
DR   WPI; [illegible]
DR   N-PSDB; AAA[illegible]
XX
PT   New human mesenchymal stem cell [illegible] polypeptides
PT   useful as reagents for chromosomal mapping, DNA fingerprinting and for
PT   detecting presence of genetic mutations or diseases affecting bone
PT   [illegible]
XX
PS   Claim [illegible]
XX
CC   New human mesenchymal stem cell cDNA's and their fragments are useful
CC   for detecting [illegible]
CC   presence of specific mRNA [illegible] type and in detecting
CC   linkage analysis, and for locating gene regions associated with
CC   genetic disease, to control gene expression. The cDNA's, fragments
CC   and expression products are useful as chromosome markers for
CC   determining the location of genes within the genome and alleles
CC   expressed, during the development of differentiated mesenchymal cell.
CC   The cDNA's are also is useful in detecting the presence of mutations
CC   in genes leading to diseases affecting bone growth, by comparing the
CC   sequences with those found in similar chromosomal locations in
CC   mammals, especially humans afflicted with diseases affecting bone
CC   growth, in human fingerprinting and in gene therapy. The polypeptides
CC   encoded by the mesenchymal stem cell cDNA's or peptide fragments,
CC   derivatives, analogs or cells expressing them are useful as
CC   immunogens for the production of antibodies
XX
SQ   Sequence   275 AA;

  Query Match             41.7%;  Score 4[illegible];  DB 21;  Length 275;
  Best Local Similarity   47.1%;  Pred. No. [illegible];
  Matches    8;  Conservative    2;  Mismatches    7;  Indels    [illegible];  Gaps    [illegible];

Qy        1 GSICVLAGMAIHFGEMP 17
              ||:|:      || || |
Db      144 qssclarhrrthtqerp 160

Search completed: August 22, 2001, 14:[illegible]
Job time: 304 sec
```

```
                      GenCore version 4.5
           Copyright (c) 1993 - 200[illegible] Compugen Ltd.

OM protein   protein search, using sw model

Run on:          August 22, 2001, 14:30:1[illegible] ; Search time 55.29 Seconds
                                                  (without alignments)
                                                  7.076 Million cell updates/sec

Title:           US-08-659-254-5
Perfect score:   103
Sequence:        1 GSTCVLAGMATHFGEMPEQ 19

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        197339 seqs, 20590346 residues

Total number of hits satisfying chosen parameters:      197339

 mum DB seq length: 0
 mum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :        Issued_Patents_AA:*
                1:  /cgn2_6/ptodata/2/[illegible]COMB.pep:*
                2:  /cgn2_6/ptodata/2/[illegible]COMB.pep:*
                3:  /cgn2_6/ptodata/2/[illegible]COMB.pep:*
                4:  /cgn2_6/ptodata/2/[illegible]COMB.pep:*
                5:  /cgn2_6/ptodata/2/[illegible]_COMB.pep:*
                6:  /cgn2_6/ptodata/2/[illegible]files1.pep:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                              SUMMARIES
                 %
Result           Query
  No.   Score    Match Length DB  ID                          Description
-------------------------------------------------------------------------------
    1     103    100.0     19  2  US-08-538-960-5             Sequence 5, Appli
    2     103    100.0     19  4  US-08-659-254-5             Sequence 5, Appli
    3     103    100.0    276  4  US-08-659-254-2             Sequence 2, Appli
    4     103    100.0    284  2  US-08-5[illegible]-2        Sequence 2, Appli
    5      89     86.4    285  4  US-08-659-254-18            Sequence 18, Appl
    6      83     80.6    285  1  US-08-4[illegible]-2        Sequence 2, Appli
    7      42     40.8    386  2  US-08-85[illegible]-1       Sequence 1, Appli
    8      42     40.8    386  3  US-09-20[illegible]284-4    Sequence 4, Appli
    9      40     38.8    390  1  US-08-121-714-6             Sequence 6, Appli
   10      40     38.8    390  1  US-08-477-[illegible]08A-6  Sequence 6, Appli
   11      40     38.8    390  2  US-08-477-[illegible]12-6   Sequence 6, Appli
   12      40     38.8    390  5  PCT-US9[illegible]08322-6   Sequence 6, Appli
   13      40     38.8    402  1  US-08-315-461-7             Sequence 7, Appli
   14      40     38.8    402  3  US-0[illegible]-2           Sequence 2, Appli
   15      40     38.8    580  2  US-0[illegible]49-12        Sequence 12, Appl
   16      40     38.8   2259  4  US-09-41[illegible]814-70   Sequence 70, Appl
   17      40     38.8   2439  3  US-0[illegible]-7           Sequence 7, Appli
   18    39.5     38.3    441  4  US-0[illegible]-764-[illegible]70-9  Sequence 9, Appli
   19      39     37.9     62  1  US-0[illegible]66A-13       Sequence 13, Appl
   20      39     37.9     62  1  US-08-683-[illegible]75-13  Sequence 13, Appl
   21      39     37.9     62  3  US-0[illegible]-13          Sequence 13, Appl
   22      39     37.9    907  3  US-0[illegible]140-4        Sequence 4, Appli
   23      39     37.9    940  4  US-08-8[illegible]12-7      Sequence 7, Appli
   24      39     37.9   1466  6  5256642-5                   Patent No. 5256642
   25      39     37.9   1466  6  5472939-5                   Patent No. 5472939
   26      39     37.9   1537  6  5256642-5                   Patent No. 5256642
   27      39     37.9   1537  6  5472939-5                   Patent No. 5472939
   28      39     37.9   1847  6  525664[illegible]           Patent No. 5256642
   29      39     37.9   1847  6  5472939-1[illegible]        Patent No. 5472939
   30      39     37.9   2039  6  5256642-2                   Patent No. 5256642
   31      39     37.9   2039  6  5472939-2                   Patent No. 5472939
   32      38     36.9    259  6  5223425-2                   Patent No. 5223425
   33      38     36.9    260  6  5223425-1                   Patent No. 5223425
   34    37.5     36.4    475  2  US-08-[illegible]-2         Sequence 2, Appli
   35    37.5     36.4    475  3  US-08-[illegible]-2         Sequence 2, Appli
   36    37.5     36.4    475  4  US-08-[illegible]87-10      Sequence 10, Appl
   37      37     35.9    170  1  US-08-2[illegible]-1        Sequence 1, Appli
   38      37     35.9    170  6  5194425-[illegible]         Patent No. 5194425
   39      37     35.9    238  4  US-08-8[illegible]-28       Sequence 28, Appl
   40      37     35.9    239  4  US-08-8[illegible]-27       Sequence 27, Appl
   41      37     35.9    607  4  US-08-[illegible]683A-59    Sequence 59, Appl
   42      37     35.9   2123  4  US-08-968-685A-10           Sequence 10, Appl
   43      37     35.9   3798  3  US-0[illegible]-5          Sequence 5, Appli
   44      37     35.9   7257  3  US-0[illegible]409-5       Sequence 5, Appli
   45    36.5     35.4    221  4  US-0[illegible]-29         Sequence 29, Appl

                              ALIGNMENTS

RESULT  1
US-08-538-960-5
; Sequence 5, Application US/08538960
; Patent No. 5872240
;  GENERAL INFORMATION:
;    APPLICANT:  Stocco, Douglas M.
;    APPLICANT:  Clark, Barbara J.
;    TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR
;    TITLE OF INVENTION:  REGULATION OF STEROIDOGENESIS
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Akin, Gump, Strauss, Hauer & Feld, L.L.P.
;      STREET:  1900 Frost Bank Plaza, 816 Congress Avenue
;      CITY:  Austin
;      STATE:  TX
;      COUNTRY:  U.S.A.
;      ZIP:  78701
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/538,960
;      FILING DATE:  Concurrently Herewith
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Mayfield, Denise L.
;      REGISTRATION NUMBER:  33,7[illegible]
;      REFERENCE/DOCKET NUMBER:  4[illegible]/DLM
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  512/499-6200
;      TELEFAX:  512/499-6290
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  19 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-538-960-5


  Query Match             100.0%;  Score 103;  DB 2;  Length 19;
  Best Local Similarity   100.0%;  Pred. No. 7e-11;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GSTCVLAGMATHFGEMPEQ 19
              |||||||||||||||||||
Db          1 GSTCVLAGMATHFGEMPEQ 19
```

```
RESULT      2
US-08-659-254-5
; Sequence 5, Application US/08659254
; Patent No. 6194559
;  GENERAL INFORMATION:
;     APPLICANT: Stocco, Douglas M
;     APPLICANT: Clark, Dr. Barbara
;     TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR REGULATION
;     TITLE OF INVENTION: OF STEROIDOGENESIS
;     NUMBER OF SEQUENCES: 1[illegible]
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE: Akin, Gump, Strauss, Hauer & Feld, L.L.P.
;       STREET: 1900 Frost Bank Plaza, 816 Congress Avenue
;       CITY: Austin
;       STATE: TX
;       COUNTRY: U.S.A.
;       ZIP: 78701
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE: Floppy disk
;       COMPUTER: IBM PC compatible
;       OPERATING SYSTEM: PC-DOS/MS-DOS
;       SOFTWARE: PatentIn Release #1.0, Version #[illegible]
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER: US/08/659,254
;       FILING DATE: [illegible]-JUN-1996
;       CLASSIFICATION: 5[illegible]
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER: US 08/[illegible]
;       FILING DATE: [illegible]
;     ATTORNEY/AGENT INFORMATION:
;       NAME: Mayfield, Denise L.
;       REGISTRATION NUMBER: [illegible]
;       REFERENCE/DOCKET NUMBER: [illegible]
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE: 512/499-6200
;       TELEFAX: 512/[illegible]
;  INFORMATION FOR SEQ ID NO: 5:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH: 1[illegible] amino acids
;       TYPE: amino acid
;       STRANDEDNESS: single
;       TOPOLOGY: linear
;     MOLECULE TYPE: peptide
US-08-659-254-5

  Query Match                [illegible]
  Best Local Similarity      [illegible]  Pred. No. [illegible]
  Matches [illegible]  Conservative [illegible]  Mismatches [illegible]  Indels [illegible]  Gaps [illegible]

Qy          1 GSICVLAGMAIHPGEMEQ [illegible]
              |||||||||||||||||
Db          1 GSICVLAGMAIHPGEMEQ [illegible]


RESULT      3
US-08-659-254-2
; Sequence 2, Application US/08659254
; Patent No. 6194559
;  GENERAL INFORMATION:
;     APPLICANT: Stocco, Douglas M
;     APPLICANT: Clark, Dr. Barbara
;     TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR REGULATION
;     TITLE OF INVENTION: OF STEROIDOGENESIS
;     NUMBER OF SEQUENCES: [illegible]
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE: Akin, Gump, Strauss, Hauer & Feld, L.L.P.
;       STREET: 1900 Frost Bank Plaza, 816 Congress Avenue
;       CITY: Austin
;       STATE: TX
;       COUNTRY: U.S.A.
;       ZIP: 78701
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE: Floppy disk
;       COMPUTER: IBM PC compatible
;       OPERATING SYSTEM: PC-DOS/MS-DOS
;       SOFTWARE: PatentIn Release #1.0, Version #[illegible]
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER: US/08/659,254
;       FILING DATE: [illegible]-JUN-1996
;       CLASSIFICATION: [illegible]
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER: [illegible]
;       FILING DATE: [illegible]
;     ATTORNEY/AGENT INFORMATION:
;       NAME: Mayfield, Denise L.
;       REGISTRATION NUMBER: [illegible]
;       REFERENCE/DOCKET NUMBER: [illegible]
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE: 512/[illegible]
;       TELEFAX: 512/[illegible]
;  INFORMATION FOR SEQ ID NO: 2:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH: 2[illegible] amino acids
;       TYPE: amino acid
;       STRANDEDNESS: single
;       TOPOLOGY: linear
;     MOLECULE TYPE: protein
US-08-659-254-2

  Query Match                [illegible]
  Best Local Similarity      [illegible]  Pred. No. [illegible]
  Matches [illegible]  Conservative [illegible]  Mismatches [illegible]  Indels [illegible]  Gaps [illegible]

Qy          1 GSICVLAGMAIHPGEMEQ [illegible]
              [illegible]
Db        [illegible] GSICVLAGMAIHPGEMEQ [illegible]


RESULT      4
US-08-[illegible]
; Sequence 2, Application US/[illegible]
; Patent No. [illegible]
;  GENERAL INFORMATION:
;     APPLICANT: Stocco, Douglas M.
;     APPLICANT: Clark, Barbara J.
;     TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
;     TITLE OF INVENTION: REGULATION OF STEROIDOGENESIS
;     NUMBER OF SEQUENCES: [illegible]
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE: Akin, Gump, Strauss, Hauer & Feld, L.L.P.
;       STREET: 1900 Frost Bank Plaza, 816 Congress Avenue
;       CITY: Austin
;       STATE: TX
;       COUNTRY: U.S.A.
;       ZIP: [illegible]
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE: Floppy disk
;       COMPUTER: IBM PC compatible
;       OPERATING SYSTEM: PC-DOS/MS-DOS
;       SOFTWARE: PatentIn Release #1.0, Version #1.[illegible]
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER: [illegible]
;       FILING DATE: Concurrently herewith
;       CLASSIFICATION: [illegible]
;     ATTORNEY/AGENT INFORMATION:
;       NAME: Mayfield, Denise L.
;       REGISTRATION NUMBER: [illegible]
;       REFERENCE/DOCKET NUMBER: [illegible]
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE: 512/[illegible]
;       TELEFAX: 512/[illegible]
;  INFORMATION FOR SEQ ID NO: 2:
```

```
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  284 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-538-960-2


  Query Match             100.0%;  Score 103;  DB 2;  Length 284;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-09;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GSTCVLAGMATHFGEMPEQ 19
              |||||||||||||||||||
Db        193 GSTCVLAGMATHFGEMPEQ 211


RESULT    5
US-08-659-254-18
; Sequence 18, Application US/08659254
; Patent No. 6194555
;  GENERAL INFORMATION:
;     APPLICANT:  Stocco, Douglas M.
;     APPLICANT:  Clark, Dr. Barbara J.
;     TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR REGULATION
;     TITLE OF INVENTION:  OF STEROIDOGENESIS
;     NUMBER OF SEQUENCES:  19
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Akin, Gump, Strauss, Hauer & Feld, L.L.P.
;       STREET:  1900 Frost Bank Plaza, 816 Congress Avenue
;       CITY:  Austin
;       STATE:  TX
;       COUNTRY:  U.S.A.
;       ZIP:  78701
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/659,254
;       FILING DATE:  07-JUN-1996
;       CLASSIFICATION:  530
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/538,960
;       FILING DATE:  04-OCT-1995
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Mayfield, Denise L.
;       REGISTRATION NUMBER:  33,732
;       REFERENCE/DOCKET NUMBER:  43375.0006
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  512/499-6200
;       TELEFAX:  512/499-6290
;   INFORMATION FOR SEQ ID NO:  18:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  285 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
US-08-659-254-18


  Query Match             86.4%;  Score 89;  DB 4;  Length 285;
  Best Local Similarity   89.5%;  Pred. No. 3.9e-07;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GSTCVLAGMATHFGEMPEQ 19
              ||||||||||| || ||||
Db        194 GSTCVLAGMATDFGNMPEQ 212


RESULT    6
US-08-410-540-2
```

```
; Sequence 2, Application US/08410540
; Patent No. 5807678
;  GENERAL INFORMATION:
;     APPLICANT:  Miller, Walter L.
;     APPLICANT:  Lin, Dong
;     APPLICANT:  Strauss III, Jerome F.
;     TITLE OF INVENTION:  IDENTIFICATION OF GENE MUTATIONS
;     TITLE OF INVENTION:  ASSOCIATED WITH CONGENITAL LIPOID ADRENAL HYPERPLASIA
;     NUMBER OF SEQUENCES:  30
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Cooley Godward Castro Huddleson & Tatum
;       STREET:  5 Palo Alto Square
;       CITY:  Palo Alto
;       STATE:  CA
;       COUNTRY:  US
;       ZIP:  94306-2155
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/410,540
;       FILING DATE:  23-MAR-1995
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Neeley, Richard L.
;       REGISTRATION NUMBER:  30,092
;       REFERENCE/DOCKET NUMBER:  UCAL-248/00US
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  415 853 5070
;       TELEFAX:  415 857 0663
;       TELEX:  380816COOLEYPA
;   INFORMATION FOR SEQ ID NO:  2:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  285 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-410-540-2


  Query Match             80.6%;  Score 83;  DB 1;  Length 285;
  Best Local Similarity   84.2%;  Pred. No. 2.8e-06;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 GSTCVLAGMATHFGEMPEQ 19
              ||||||||| | || ||||
Db        194 GSTCVLAGMSTDFGNMPEQ 212


RESULT    7
US-08-858-052-4
; Sequence 4, Application US/08858052
; Patent No. 5849498
;  GENERAL INFORMATION:
;     APPLICANT:  Bandman, Olga
;     APPLICANT:  Guegler, Karl J.
;     APPLICANT:  Shah, Purvi
;     APPLICANT:  Corley, Neil C.
;     TITLE OF INVENTION:  HUMAN 3 HYDROXYISOBUTYRYL-COENZYME A
;     TITLE OF INVENTION:  HYDROLASE
;     NUMBER OF SEQUENCES:  4
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;       STREET:  3174 Porter Drive
;       CITY:  Palo Alto
;       STATE:  CA
;       COUNTRY:  USA
;       ZIP:  94304
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Diskette
;       COMPUTER:  IBM Compatible
;          OPERATING SYSTEM: [illegible]
;          SOFTWARE: FastSEQ for Windows Version [illegible]
;       CURRENT APPLICATION DATA:
;          APPLICATION NUMBER: [illegible]
;          FILING DATE: Herewith
;          CLASSIFICATION: [illegible]
;       PRIOR APPLICATION DATA:
;          APPLICATION NUMBER:
;          FILING DATE:
;       ATTORNEY/AGENT INFORMATION:
;          NAME: Billings, Lucy J
;          REGISTRATION NUMBER: [illegible]
;          REFERENCE/DOCKET NUMBER: [illegible]
;       TELECOMMUNICATION INFORMATION:
;          TELEPHONE: 41[illegible]
;          TELEFAX: 41[illegible]
;          TELEX:
;  INFORMATION FOR SEQ ID NO: 4:
;       SEQUENCE CHARACTERISTICS:
;          LENGTH: [illegible] amino acids
;          TYPE: amino acid
;          STRANDEDNESS: single
;          TOPOLOGY: linear
;       IMMEDIATE SOURCE:
;          LIBRARY: GenBank
;          CLONE: 485[illegible]
US-08-858-[illegible]-4

  Query Match            [illegible]
  Best Local Similarity  [illegible]  Pred. No. [illegible]
  Matches [illegible]  Conservative [illegible]  Mismatches [illegible]

Qy          1 [illegible]VLA[illegible]MA[illegible]GEM[illegible]
              [illegible]
Db        207 GA[illegible]AFHAGATHEVE[illegible]


RESULT 8
US-09-200-284-4
; Sequence 4, Application US/09200284
; Patent No. [illegible]
;  GENERAL INFORMATION:
;       APPLICANT: Bandman, Olga
;       APPLICANT: Guegler, Karl J.
;       APPLICANT: Shah, Purvi
;       APPLICANT: Corley, Neil C.
;       TITLE OF INVENTION: HUMAN [illegible]
;       TITLE OF INVENTION: [illegible]YDROLASE
;       NUMBER OF SEQUENCES: [illegible]
;       CORRESPONDENCE ADDRESS:
;          ADDRESSEE: Incyte Pharmaceuticals, Inc.
;          STREET: [illegible] Porter Drive
;          CITY: Palo Alto
;          STATE: CA
;          COUNTRY: USA
;          ZIP: 94304
;       COMPUTER READABLE FORM:
;          MEDIUM TYPE: Diskette
;          COMPUTER: IBM Compatible
;          OPERATING SYSTEM: [illegible]
;          SOFTWARE: FastSEQ for Windows Version [illegible]
;       CURRENT APPLICATION DATA:
;          APPLICATION NUMBER: [illegible]
;          FILING DATE:
;          CLASSIFICATION:
;       PRIOR APPLICATION DATA:
;          APPLICATION NUMBER: [illegible]
;          FILING DATE:
;       ATTORNEY/AGENT INFORMATION:
;          NAME: Billings, Lucy J
;          REGISTRATION NUMBER: [illegible]
;          REFERENCE/DOCKET NUMBER: [illegible]
```

```
;       TELECOMMUNICATION INFORMATION:
;          TELEPHONE: 41[illegible]
;          TELEFAX: 41[illegible]
;          TELEX:
;  INFORMATION FOR SEQ ID NO: [illegible]
;       SEQUENCE CHARACTERISTICS:
;          LENGTH: [illegible] amino acids
;          TYPE: amino acid
;          STRANDEDNESS: single
;          TOPOLOGY: linear
;       IMMEDIATE SOURCE:
;          LIBRARY: GenBank
;          CLONE: [illegible]
US-[illegible]-4

  Query Match            [illegible]  Score [illegible]  DB [illegible]  Length [illegible]
  Best Local Similarity  [illegible]  Pred. No. [illegible]
  Matches [illegible]  Conservative [illegible]  Mismatches [illegible]  Indels [illegible]  Gaps [illegible]

Qy          1 [illegible]VLA[illegible]MA[illegible]GEM[illegible]

Db        [illegible] [illegible]AFHAGATHEVE[illegible]


RESULT 9
US-[illegible]
; Sequence [illegible], Application US/[illegible]
; Patent No. [illegible]
;  GENERAL INFORMATION:
;       APPLICANT: [illegible]
;       TITLE OF INVENTION: [illegible] WITH
;       TITLE OF INVENTION: [illegible] REGULATING ACTIVITY
;       NUMBER OF SEQUENCES: [illegible]
;       CORRESPONDENCE ADDRESS:
;          ADDRESSEE: [illegible]
;          STREET: 225 Franklin Street
;          CITY: Boston
;          STATE: Massachusetts
;          COUNTRY: USA
;          ZIP: [illegible]
;       COMPUTER READABLE FORM:
;          MEDIUM TYPE: [illegible] Diskette, [illegible] MB
;          COMPUTER: IBM PS/2 Model [illegible] or 55SX
;          OPERATING SYSTEM: MS-DOS (Version [illegible])
;          SOFTWARE: WordPerfect (Version [illegible])
;       CURRENT APPLICATION DATA:
;          APPLICATION NUMBER: [illegible]
;          FILING DATE:
;          CLASSIFICATION: [illegible]
;       PRIOR APPLICATION DATA:
;          APPLICATION NUMBER: [illegible]
;          FILING DATE: [illegible]
;          APPLICATION NUMBER: [illegible]844[illegible]
;          FILING DATE: [illegible]
;          APPLICATION NUMBER: [illegible]
;          FILING DATE: [illegible]
;       ATTORNEY/AGENT INFORMATION:
;          NAME: [illegible]
;          REGISTRATION NUMBER: [illegible]
;          REFERENCE/DOCKET NUMBER: [illegible]
;       TELECOMMUNICATION INFORMATION:
;          TELEPHONE: [illegible]
;          TELEFAX: [illegible]
;          TELEX: [illegible]
;  INFORMATION FOR SEQ ID NO: [illegible]
;       SEQUENCE CHARACTERISTICS:
;          LENGTH: [illegible]
;          TYPE: [illegible]
;          STRANDEDNESS:
;          TOPOLOGY: [illegible]
[illegible]
```

```
  Query Match             38.8%;  Score 40;  DB 1;  Length 390;
  Best Local Similarity   53.8%;  Pred. No. 59;
  Matches    7;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy        3 TCVLAGMATHFGE 15
            ||:: |:|  |||
Db        8 TCLVLGLAIVFGE 20


RESULT  10
US-08-477-108A-6
; Sequence 6, Application US/08477108A
; Patent No. 5801001
;  GENERAL INFORMATION:
;    APPLICANT:  Sager, Ruth
;    APPLICANT:  Zou, Zhiqiang
;    APPLICANT:  Anisowicz, Anthony
;    TITLE OF INVENTION:  MASPIN, A NOVEL SERPIN WITH
;    TITLE OF INVENTION:  TUMOR SUPPRESSING ACTIVITY
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson P.C.
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;      COMPUTER:  IBM PS/2 Model 50Z or 55SX
;      OPERATING SYSTEM:  MS-DOS (Version 5.0)
;      SOFTWARE:  WordPerfect (Version 5.1)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/477,108A
;      FILING DATE:  June 7, 1995
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/121,714
;      FILING DATE:  09/01/93
;      APPLICATION NUMBER:  07/938,823
;      FILING DATE:  09/01/92
;      APPLICATION NUMBER:  07/844,296
;      FILING DATE:  02/28/92
;      APPLICATION NUMBER:  07/662,216
;      FILING DATE:  02/28/91
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fraser, Janis K.
;      REGISTRATION NUMBER:  34,819
;      REFERENCE/DOCKET NUMBER:  06570/002002
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 542-5070
;      TELEFAX:  (617) 542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  390
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
US-08-477-108A-6


  Query Match             38.8%;  Score 40;  DB 1;  Length 390;
  Best Local Similarity   53.8%;  Pred. No. 59;
  Matches    7;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy        3 TCVLAGMATHFGE 15
            ||:: |:|  |||
Db        8 TCLVLGLAIVFGE 20


RESULT  11
US-08-477-112-6
; Sequence 6, Application US/08477112
; Patent No. 5905023
;  GENERAL INFORMATION:
;    APPLICANT:  Sager, Ruth
;    TITLE OF INVENTION:  MASPIN, A NOVEL SERPIN WITH
;    TITLE OF INVENTION.  TUMOR SUPPRESSING ACTIVITY
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson P.C.
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
,      COMPUTER.  IBM PS/2 Model 50Z or 55SX
,      OPERATING SYSTEM:  MS-DOS (Version 5.0)
;      SOFTWARE:  WordPerfect (Version 5.1)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER.  US/08/477,112
;      FILING DATE:
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/121,714
;      FILING DATE:  09/01/93
,      APPLICATION NUMBER.  07/938,823
;      FILING DATE:  09/01/92
;      APPLICATION NUMBER:  07/844,296
       FILING DATE:  02/28/92
;      APPLICATION NUMBER:  07/662,216
;      FILING DATE:  02/28/91
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fraser, Janis K.
;      REGISTRATION NUMBER:  34,819
;      REFERENCE/DOCKET NUMBER:  06570/002003
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 542-5070
;      TELEFAX:  (617) 542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  390
;      TYPE:  amino acid
,      STRANDEDNESS:
;      TOPOLOGY:  linear
US-08-477-112-6


  Query Match             38.8%;  Score 40;  DB 2;  Length 390;
  Best Local Similarity   53.8%;  Pred. No. 59;
  Matches    7;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy        3 TCVLAGMATHFGE 15
            ||:: |:|  |||
Db        8 TCLVLGLAIVFGE 20


RESULT  12
PCT-US93-08322-6
; Sequence 6, Application PC/TUS9308322
;  GENERAL INFORMATION:
;    APPLICANT:  Sager, Ruth
,    TITLE OF INVENTION.  MASPIN, A NOVEL SERPIN WITH TUMOR SUPPRESSING ACTIVITY
,    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
;      COMPUTER: IBM PS/2 Model [illegible]
;      OPERATING SYSTEM: MS-DOS (Version [illegible])
;      SOFTWARE: WordPerfect (Version [illegible])
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER: PCT/US93/[illegible]
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER: [illegible]
;      FILING DATE: 09/[illegible]/92
;      APPLICATION NUMBER: [illegible]/841,[illegible]
;      FILING DATE: 02/[illegible]/92
;      APPLICATION NUMBER: [illegible]
;      FILING DATE: 02/[illegible]/91
;    ATTORNEY/AGENT INFORMATION:
;      NAME: Fraser, [illegible] K
;      REGISTRATION NUMBER: [illegible]
;      REFERENCE/DOCKET NUMBER: [illegible]
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE: ([illegible]) [illegible]
;      TELEFAX: ([illegible]) [illegible]
;      TELEX: [illegible]
;  INFORMATION FOR SEQ ID NO: [illegible]:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH: [illegible]
;      TYPE: amino acid
;      STRANDEDNESS:
;      TOPOLOGY: linear
PCT-US93-08[illegible]2-6

  Query Match             88.8%;  Score 1[illegible];  DB [illegible];  Length [illegible];
  Best Local Similarity   58.8%;  Pred. No. [illegible];
  Matches    7;  Conservative    3;  Mismatches  [illegible];  Indels  [illegible];  Gaps  [illegible];

Qy          3 ICVLAGMATHE[illegible] 1
              |:::|:|
Db          8 ICVLGLAIVE[illegible] 2


RESULT 13
US-08-335-461-7
; Sequence 7, Application US/08335461
; Patent No. 5639726
; GENERAL INFORMATION:
;    APPLICANT: Lawrence, Daniel A
;    APPLICANT: Ginsburg, David
;    APPLICANT: Shore, Joseph D
;    APPLICANT: Fay, William P
;    APPLICANT: [illegible]
;    APPLICANT: Francis-Chmura, Anne Marie
;    APPLICANT: [illegible]
;    APPLICANT: [illegible]
;    TITLE OF INVENTION: Peptide Mediated Enhancement [illegible]
;    TITLE OF INVENTION: Thrombolysis Methods and Compositions
;    NUMBER OF SEQUENCES: [illegible]
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE: Arnold, White & Durkee
;      STREET: P.O. Box 4433
;      CITY: Houston
;      STATE: TX
;      COUNTRY: USA
;      ZIP: 77210
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE: Floppy disk
;      COMPUTER: IBM PC compatible
;      OPERATING SYSTEM: PC-DOS/MS-DOS
;      SOFTWARE: PatentIn Release #1.0, Version #1.[illegible]
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER: US/08/335,461
;      FILING DATE: [illegible]
```

```
;      CLASSIFICATION: [illegible]
;    ATTORNEY/AGENT INFORMATION:
;      NAME: Parker, David L
;      REGISTRATION NUMBER: [illegible]
;      REFERENCE/DOCKET NUMBER: UMI[illegible]
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE: ([illegible]) 4[illegible]
;      TELEFAX: ([illegible]) [illegible]
;  INFORMATION FOR SEQ ID NO: [illegible]:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH: 4[illegible] amino acids
;      TYPE: amino acid
;      STRANDEDNESS: [illegible]
;      TOPOLOGY: linear
;    MOLECULE TYPE: [illegible]
US-08-335-461-[illegible]

  Query Match             [illegible]%;  Score 4;  DB [illegible];  Length 4[illegible];
  Best Local Similarity   [illegible]%;  Pred. No. [illegible];
  Matches  [illegible];  Conservative  [illegible];  Mismatches  [illegible];  Indels  [illegible];  Gaps  [illegible];

Qy          [illegible] VLAGMATHE [illegible]
              [illegible]
Db          [illegible] VLGLAIVE [illegible]


RESULT [illegible]
US-08-[illegible]
; Sequence [illegible], Application US/[illegible]
; Patent No. [illegible]
; GENERAL INFORMATION:
;    APPLICANT: LAWRENCE, DANIEL A
;    APPLICANT: STEFANSSON, STEINGRIMUR [illegible]
;    TITLE OF INVENTION: MUTANT PLASMINOGEN ACTIVATOR INHIBITOR
;    TITLE OF INVENTION: TYPE 1 (PAI-1) AND USES THEREOF
;    NUMBER OF SEQUENCES: [illegible]
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE: MORRISON & FOERSTER
;      STREET: 2000 PENNSYLVANIA AVENUE, NW
;      CITY: WASHINGTON
;      STATE: DC
;      COUNTRY: USA
;      ZIP: [illegible]
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE: Floppy disk
;      COMPUTER: IBM PC compatible
;      OPERATING SYSTEM: PC-DOS/MS-DOS
;      SOFTWARE: PatentIn Release #1.0, Version #1.[illegible]
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER: [illegible]
;      FILING DATE:
;      CLASSIFICATION: [illegible]
;    ATTORNEY/AGENT INFORMATION:
;      NAME: [illegible], SAMUEL
;      REGISTRATION NUMBER: [illegible]
;      REFERENCE/DOCKET NUMBER: [illegible]
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE: (202) 887-[illegible]
;      TELEFAX: (202) [illegible]
;      TELEX: [illegible] MRSNFOERSWSH
;  INFORMATION FOR SEQ ID NO: [illegible]:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH: [illegible] amino acids
;      TYPE: amino acid
;      TOPOLOGY: linear
;    MOLECULE TYPE: protein
US-08-[illegible]

  Query Match             [illegible]%;  Score 4;  DB [illegible];  Length 4[illegible];
  Best Local Similarity   [illegible]%;  Pred. No. [illegible];
  Matches  [illegible];  Conservative  [illegible];  Mismatches  [illegible];  Indels  [illegible];  Gaps  [illegible];
```

```
Qy       3 TCVLACMATHFGE 15
           ||::  |:|   |||
Db       8 TCLVLGLALVFGE 20


RESULT  15
US-08-677-049-12
; Sequence 12, Application US/08677049
; Patent No. 5858707
;  GENERAL INFORMATION:
;    APPLICANT:  Guimaraes, M. Jorge
;    APPLICANT:  Bazan, J. Fernando
;    APPLICANT:  McClanahan, Terrill K.
;    APPLICANT:  Zlotnik, Albert
;    TITLE OF INVENTION:  PURIFIED MAMMALIAN NUCLEOBASE PERMEASES;
;    TITLE OF INVENTION:  NUCLEIC ACIDS; ANTIBODIES
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  DNAX Research Institute
;      STREET:  901 California Avenue
;      CITY:  Palo Alto
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94304-1104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/677,049
;      FILING DATE:  03-JUL-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/000,788
;      FILING DATE:  03-JUL-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ching, Edwin P.
;      REGISTRATION NUMBER:  34,090
;      REFERENCE/DOCKET NUMBER:  DX0511
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-852-9196
;      TELEFAX:  415-496-1200
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  580 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    FEATURE:
;      NAME/KEY:  Region
;      LOCATION:  207..236
;      OTHER INFORMATION:  /note  "Encompasses IM 4 of Figure
;      OTHER INFORMATION:  4"
;    FEATURE:
;      NAME/KEY:  Region
;      LOCATION:  437..499
;      OTHER INFORMATION:  /note= "Encompasses IM 9 and IM 10
;      OTHER INFORMATION:  of Figure 4"
US-08-677-049-12

  Query Match             38.8%;  Score 4[illegible];  DB 2;  Length 580;
  Best Local Similarity   63.6%;  Pred. No. 78;
  Matches    7;  Conservative    2;  Mismatches    2;  Indels    [illegible];  Gaps    [illegible];

Qy       1 GSTCVLAGMAT 11
           | ||:|||: |
Db     388 GITCLLAGLCT 398
```

```
Search completed: August 22, 2001, 14:3[illegible]
Job time: 365 sec
```

```
                    GenCore version 4.5
          Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on:         August 22, 2001, 14:30:54 ; Search time 35.8 Seconds
                                            (without alignments)
                                            11.482 Million cell updates/sec

Title:          US-08-659-254-4
Perfect score:  66
Sequence:       1 ALGILNNQEGWK 12

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters:      93435

   mum DB seq length: 0
   mum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :      SwissProt_39:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                               SUMMARIES
            %
Result      Query
  No. Score Match Length DB ID                          Description
------------------------------------------------------------------------
   1    66  100.0   284  1  STAR_MOUSE                  P51557 mus musculus
   2    66  100.0   284  1  STAR_RAT                    P97826 rattus norv
   3    61   92.4   285  1  STAR_HUMAN                  P49675 homo sapien
   4    60   90.9   285  1  STAR_HORSE                  O46689 equus cabal
   5    55   83.3   284  1  STAR_MESAU                  P70114 mesocricetu
   6    55   83.3   285  1  STAR_BOVIN                  Q28918 bos taurus
   7    54   81.8   285  1  STAR_PIG                    Q28996 sus scrofa
   8    47   71.2   157  1  UBC3_SCHM                   P40984 schistosoma
   9    43   65.2   800  1  GUN_BACSI                   P06564 bacillus sp
  10    39   59.1  1742  1  GUNA_CALSA                  P22534 caldocellum
  11    38   57.6   184  1  KPTA_ECOLI                  P39380 escherichia
  12    38   57.6   469  1  BGL1_BACSU                  P40740 bacillus su
  13    38   57.6   475  1  TBE_HUMAN                   Q9UJT0 homo sapien
  14    38   57.6   620  1  YG18_YEAST                  P53207 saccharomyc
  15    38   57.6   825  1  GUN3_BACSM                  P19570 bacillus sp
  16    38   57.6   865  1  CYAA_PROMI                  Q59685 proteus mir
  17    37   56.1   359  1  MAS2_AGRRI                  P27873 agrobacteri
  18    37   56.1   858  1  CNRA_BOVIN                  P11541 bos taurus
  19    37   56.1   858  1  CNRA_MOUSE                  P27664 mus musculu
  20    37   56.1   859  1  CNRA_HUMAN                  P16499 homo sapien
  21    37   56.1   860  1  CNRA_CANFA                  Q28293 canis famil
  22    37   56.1  1511  1  PDR5_YEAST                  P33302 saccharomyc
  23    37   56.1  1529  1  PDRF_YEAST                  Q04182 saccharomyc
  24    36   54.5   322  1  GRP2_MOUSE                  Q9QY17 m grb2 rela
  25    36   54.5   330  1  GRP2_HUMAN                  O75791 h grb2 rela
  26    36   54.5   333  1  A85C_MYCLE                  Q05862 mycobacteri
  27    36   54.5   340  1  A85C_MYCTU                  P31953 mycobacteri
  28    36   54.5   349  1  WN7B_MOUSE                  P28047 mus musculu
  29    36   54.5   394  1  EFTU_BACFR                  P33165 bacteroides
  30    36   54.5   394  1  EFTU_MYCGA                  P18906 mycoplasma
  31    36   54.5   395  1  EFTU_CYTLY                  P42471 cytophaga l
  32    36   54.5   458  1  [illegible]                 [illegible]
  33    36   54.5  1256  1  ATL_STAAU                   P52081 staphylococ
  34    36   54.5  1501  1  CDR1_CANAL                  O42690 candida alb
  35    35   53.0   274  1  [illegible]                 Q58[illegible] methanococ
  36    35   53.0   294  1  RIP1_TRIAN                  P56626 trichosanth
  37    35   53.0   358  1  25A1_RAT                    Q05961 rattus norv
  38    35   53.0   358  1  OAS1_RAT                    Q05961 rattus norv
  39    35   53.0   368  1  LYS1_SCHPO                  Q09694 schizosacch
  40    35   53.0   393  1  DOM3_CAEEL                  Q10[illegible] caenorhabdi
  41    35   53.0   406  1  TRKA_METMA                  P39449 methanosarc
  42    35   53.0   439  1  CKI3_SCHPO                  O74135 schizosacch
  43    35   53.0   446  1  CKI1_SCHPO                  P40233 schizosacch
  44    35   53.0   486  1  YN8  YEAST                  P53742 saccharomyc
  45    35   53.0   520  1  YZ08_METJA                  Q60270 methanococc

                               ALIGNMENTS

RESULT    1
STAR_MOUSE
ID   STAR_MOUSE     STANDARD;      PRT;   284 AA.
AC   P51557;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN PRECURSOR (LUTEINIZING HORMONE
DE   INDUCED PROTEIN)
GN   STAR.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX   MEDLINE=95050616; PubMed=7961773;
RA   Clark B.J., Wells J., King S.R., Stocco D.M.;
RT   "The purification, cloning, and expression of a novel luteinizing
RT   hormone-induced mitochondrial protein in MA-10 mouse Leydig tumor
RT   cells. Characterization of the steroidogenic acute regulatory protein
RT   (StAR).";
RL   J. Biol. Chem. 269:28314-28322(1994).
CC   -!- FUNCTION: PLAYS A KEY ROLE IN STEROID HORMONE SYNTHESIS BY
CC       ENHANCING THE METABOLISM OF CHOLESTEROL INTO PREGNENOLONE.
CC       TRANSFERASE THAT BINDS TO AND TRANSFERS CHOLESTEROL THROUGH THE
CC       INTERMEMBRANE SPACE OF THE MITOCHONDRION (PROBABLE).
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL.
CC   -!- SIMILARITY: BELONGS TO THE STAR FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; L36062; AAB94783.1; -.
DR   MGD; MGI:102760; Star.
DR   InterPro; IPR000799; -.
DR   InterPro; IPR002913; -.
DR   Pfam; PF01852; START; 1.
DR   PRINTS; PR00978; STARPROTEIN.
KW   Lipid-binding; Lipid transport; Transit peptide; Steroidogenesis;
KW   Mitochondrion; Transit peptide.
FT   TRANSIT       1      ?       MITOCHONDRION (POTENTIAL).
FT   CHAIN         ?    284       STEROIDOGENIC ACUTE REGULATORY PROTEIN.
FT   DOMAIN      173    174       POLY-ALA.
SQ   SEQUENCE   284 AA;  31626 MW;  B81E52144B9B72 ?E?64;

  Query Match             100.0%;  Score 66;  DB 1;  Length 284;
  Best Local Similarity   100.0%;  Pred. No. 2.5e-05;
  Matches   12;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          1 ALGILNNQ[illegible]W[illegible] 1[illegible]
              |||||||| | |
Db         8[illegible] ALGILNNQ[illegible]W[illegible] 9[illegible]


RESULT   2
STAR_RAT
ID   STAR_RAT       STANDARD;      PRT;   [illegible] AA.
AC   P97826; P97628;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN [illegible]
GN   STAR.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; [illegible]
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=WISTAR; TISSUE=[illegible];
RX   MEDLINE=974[illegible]; PubMed=93[illegible];
RA   Mizutani T., Sonoda Y., Minegishi T., Wakabayashi K., M[illegible];
RT   "Molecular cloning, characterization and [illegible]
RT   steroidogenic acute regulatory protein (StAR) [illegible]."
RL   Life Sci. 61:[illegible](199[illegible]).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY; TISSUE=[illegible];
RX   MEDLINE=9716[illegible]; PubMed=90[illegible];
RA   Lee H.K., Ahn R.[illegible], Kwon H.B.[illegible];
RT   "Nucleotide sequence of rat steroidogenic acute regulatory protein
RT   complementary DNA.";
RL   Biochem. Biophys. Res. Commun. [illegible]
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY; TISSUE=Adrenal gland;
RX   MEDLINE=9818[illegible]; PubMed=9[illegible];
RA   Ariyoshi N., Kim Y.[illegible], Artemenko I.[illegible],
RA   Jefcoate C.R.;
RT   "Characterization of the rat Star gene [illegible]
RT   3.5 kilobase pair mRNA: ACTH stimulation [illegible]
RT   precedes elevation of Star mRNA and protein."
RL   J. Biol. Chem. [illegible](199[illegible]).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY; TISSUE=Adrenal gland;
RX   MEDLINE=99048[illegible]; PubMed=98[illegible];
RA   Furukawa A., Miyatake A., [illegible];
RT   "Steroidogenic acute regulatory protein (StAR) [illegible]
RT   constitutively expressed in the adult rat [illegible]
RT   colocalization of StAR, cytochrome P-450scc (CYP11A1), [illegible]
RT   hydroxysteroid dehydrogenase in the rat brain."
RL   J. Neurochem. 71:[illegible](199[illegible]).
CC   -!- FUNCTION: PLAYS A KEY ROLE IN STEROID HORMONE [illegible]
CC       ENHANCING THE METABOLISM OF [illegible]
CC       TRANSPORTER THAT BINDS [illegible] AND TRANSFER[illegible]
CC       INTERMEMBRANE SPACE OF THE MITOCHONDRION [illegible].
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIA.
CC   -!- SIMILARITY: BELONGS TO THE STAR FAMILY.
CC
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC
DR   EMBL; AB001349; BAA19[illegible]; [illegible].
DR   EMBL; U76419; AAB[illegible]; [illegible].
DR   EMBL; AF044081; AAC[illegible]; [illegible].
DR   EMBL; AB006007; BAA3[illegible]; [illegible].
DR   [illegible]
DR   [illegible]
DR   [illegible]
DR   [illegible] START[illegible]
KW   [illegible] transport; Transport; [illegible]
KW   Mitochondrion; Transit peptide.
FT   TRANSIT       1     [illegible]       MITOCHONDRION (POTENTIAL).
FT   CHAIN         [illegible]   284       STEROIDOGENIC ACUTE REGULATORY PROTEIN.
FT   CONFLICT      1     [illegible]       MLLATFK[illegible] -> [illegible]LWESTAVLPPLYVEPPPPSMPPV[illegible]N
FT                                         VYP[illegible]VVVNBY[illegible]AF[illegible]
FT                                         [illegible]VLAYE[illegible] (IN REF. 2).
SQ   SEQUENCE   284 AA;  [illegible] MW;  [illegible] CRC64;

  Query Match             [illegible];  Score [illegible];  DB [illegible];  Length 284;
  Best Local Similarity   [illegible];  Pred. No. [illegible];
  Matches [illegible];  Conservative [illegible];  Mismatches [illegible];  Indels [illegible];  Gaps [illegible];

Qy          1 [illegible]NNQ[illegible]WK [illegible]

Db         [illegible] NNQ[illegible]WK [illegible]


RESULT   3
STAR_HUMAN
ID   STAR_HUMAN     STANDARD;      PRT;   [illegible] AA.
AC   [illegible]
DT   [illegible]-FEB-[illegible] (Rel. [illegible], Created)
DT   [illegible]-FEB-[illegible] (Rel. [illegible], Last sequence update)
DT   [illegible]-MAY-[illegible] (Rel. [illegible], Last annotation update)
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN [illegible] PRECURSOR.
GN   STAR.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Adrenal cortex;
RX   MEDLINE=[illegible]; PubMed=[illegible];
RA   [illegible] III, Lin D.,
RA   Miller W.L., Patterson D., Clancy K.P., Hart I.M., Clark B.J.,
RA   [illegible];
RT   "Human steroidogenic acute regulatory protein: functional activity in
RT   [illegible] expression, and mapping of the
RT   structural gene to [illegible] and a pseudogene to chromosome [illegible]."
RL   Proc. Natl. Acad. Sci. U.S.A. [illegible]:4778-4782(199[illegible]).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=[illegible];
RX   MEDLINE=[illegible]; PubMed=[illegible];
RA   Sugawara T., [illegible] A., Martin [illegible], Javitt N.B., Miller W.L.,
RA   [illegible];
RT   "Structure of the human steroidogenic acute regulatory protein (StAR)
RT   gene: StAR [illegible] hydroxylase
RT   activity."
RL   Biochemistry [illegible](199[illegible]).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=[illegible]; PubMed=[illegible];
RA   [illegible] A., [illegible]dwards [illegible], M[illegible] H.M., [illegible];
RA   Lefebvre [illegible];
RT   "The human steroidogenic acute regulatory (StAR) gene is expressed in
RT   the urogenital system and encodes a mitochondrial polypeptide.";
RL   Biochim. Biophys. Acta [illegible](199[illegible]).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   TISSUE=[illegible];
RA   Yu W., [illegible], [illegible] A.;
RL   Submitted (NOV-[illegible]) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   VARIANT[illegible]
RX   MEDLINE=[illegible]; PubMed=[illegible];
```

```
RA   Bose H.S., Sugawara T., Strauss J.F. III, Miller W.L.;
RT   "The pathophysiology and genetics of congenital lipoid adrenal
RT   hyperplasia.";
RL   New Engl. J. Med. 335:1870-1878(1996).
RN   [6]
RP   VARIANT ALA-203.
RX   MEDLINE=98112495; PubMed=9452116;
RA   Katsumata N., Tanae A., Shinagawa T., Nagashima-Miyokawa A.,
RA   Shimizu M., Yasunaga T., Tanaka T., Hibi I.;
RT   "A novel frameshift mutation 840delA and a novel polymorphism D203A in
RT   the steroidogenic acute regulatory protein gene in a Japanese patient
RT   with congenital lipoid adrenal hyperplasia.";
RL   Hum. Mutat. Suppl. 1:S304-S307(1998).
CC   -!- FUNCTION: PLAYS A KEY ROLE IN STEROID HORMONE SYNTHESIS BY
CC       ENHANCING THE METABOLISM OF CHOLESTEROL INTO PREGNENOLONE.
CC       TRANSPORTER THAT BINDS TO AND TRANSPORT CHOLESTEROL THROUGH THE
CC       INTERMEMBRANE SPACE OF THE MITOCHONDRION (PROBABLE).
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN GONADS, ADRENAL CORTEX AND
CC       KIDNEY.
CC   -!- DISEASE: DEFECTS IN STAR CAUSE CONGENITAL LIPOID ADRENAL
CC       HYPERPLASIA (CLAH), THE MOST SEVERE FORM OF ADRENAL HYPERPLASIA.
CC   -!- SIMILARITY: BELONGS TO THE STAR FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; U17280; AAC50141.1; -.
DR   EMBL; U29105; AAC50234.1; -.
DR   EMBL; U29099; AAC50234.1; JOINED.
DR   EMBL; U29100; AAC50234.1; JOINED.
DR   EMBL; U29101; AAC50234.1; JOINED.
DR   EMBL; U29102; AAC50234.1; JOINED.
DR   EMBL; U29103; AAC50234.1; JOINED.
DR   EMBL; U29104; AAC50234.1; JOINED.
DR   EMBL; S79669; AAB35726.1; -.
DR   EMBL; AF035277; AAB88174.1; -.
DR   MIM; 600617; -.
DR   MIM; 201710; -.
DR   InterPro; IPR000799; -.
DR   InterPro; IPR002913; -.
DR   Pfam; PF01852; START; 1.
DR   PRINTS; PR00978; STARPROTEIN.
KW   Lipid-binding; Lipid transport; Transport; Steroidogenesis;
KW   Mitochondrion; Transit peptide; Disease mutation; Polymorphism.
FT   TRANSIT       1      ?       MITOCHONDRION (POTENTIAL).
FT   CHAIN         ?    285       STEROIDOGENIC ACUTE REGULATORY PROTEIN.
FT   VARIANT     182    182       R -> L (IN CLAH).
FT                                /FTId=VAR_005627.
FT   VARIANT     203    203       D -> A.
FT                                /FTId=VAR_005628.
SQ   SEQUENCE   285 AA;  31958 MW;  D3830EDC266838AD CRC64;

  Query Match             92.4%;  Score 6[illegible];  DB 1;  Length 285;
  Best Local Similarity   91.7%;  Pred. No. 0.00077;
  Matches   11;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ALGILNNQEGWK 12
              |||||:||||||
Db         86 ALGILSNQEGWK 97


RESULT   4
STAR_HORSE
ID   STAR_HORSE     STANDARD;      PRT;   285 AA.
AC   O46689;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN PRECURSOR.
GN   STAR.
OS   Equus caballus (Horse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX   NCBI_TaxID 9796;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Follicle;
RX   MEDLINE=99124192; PubMed=9927294;
RA   Kerban A., Boerboom D., Sirois J.;
RT   "Human chorionic gonadotropin induces an inverse regulation of
RT   steroidogenic acute regulatory protein messenger ribonucleic acid in
RT   theca interna and granulosa cells of equine preovulatory follicles.";
RL   Endocrinology 140:667-674(1999).
CC   -!- FUNCTION: PLAYS A KEY ROLE IN STEROID HORMONE SYNTHESIS BY
CC       ENHANCING THE METABOLISM OF CHOLESTEROL INTO PREGNENOLONE.
CC       TRANSPORTER THAT BINDS TO AND TRANSPORT CHOLESTEROL THROUGH THE
CC       INTERMEMBRANE SPACE OF THE MITOCHONDRION (PROBABLE).
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL.
CC   -!- SIMILARITY: BELONGS TO THE STAR FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AF031697; AAC04704.1; -.
DR   EMBL; AF031696; AAC04703.1; -.
DR   InterPro; IPR000799; -.
DR   InterPro; IPR002913; -.
DR   Pfam; PF01852; START; 1.
DR   PRINTS; PR00978; STARPROTEIN.
KW   Lipid-binding; Lipid transport; Transport; Steroidogenesis;
KW   Mitochondrion; Transit peptide.
FT   TRANSIT       1      ?       MITOCHONDRION (POTENTIAL).
FT   CHAIN         ?    285       STEROIDOGENIC ACUTE REGULATORY PROTEIN.
SQ   SEQUENCE   285 AA;  31852 MW;  [illegible] CRC64;

  Query Match             90.9%;  Score [illegible];  DB 1;  Length 285;
  Best Local Similarity   91.7%;  Pred. No. 0.12;
  Matches   11;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 ALGILNNQEGWK 12
              ||||| ||||||
Db         86 ALGILRNQEGWK 97


RESULT   5
STAR_MESAU
ID   STAR_MESAU     STANDARD;      PRT;   284 AA.
AC   P70114;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN PRECURSOR.
GN   STAR.
OS   Mesocricetus auratus (Golden hamster).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC   Mesocricetus.
OX   NCBI_TaxID 10036;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97124765; PubMed=8969935;
RA   Fleury A., Cloutier M., Ducharme L., Lefebvre A., Lehoux J.,
RA   Lehoux J.-G.;
```

```
RT   "Adrenocorticotropin regulates the [illegible]
RT   regulatory (StAR) protein mRNA in hamster adrenal [illegible]
RL   Endocr. Res. [illegible](1996).
CC   -!- FUNCTION: PLAYS A KEY ROLE IN STEROID HORMONE SYNTHESIS BY
CC       ENHANCING THE METABOLISM OF [illegible]
CC       TRANSPORTER THAT BINDS TO AND TRANSPORT [illegible]
CC       INTERMEMBRANE SPACE OF THE MITOCHONDRION [illegible]
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL.
CC   -!- SIMILARITY: BELONGS TO THE STAR FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced [illegible]
CC   between the Swiss Institute of Bioinformatics [illegible]
CC   the European Bioinformatics Institute. There are [illegible]
CC   use by non-profit institutions as long as [illegible]
CC   modified and this statement is not removed. [illegible]
CC   entities requires a license agreement (See [illegible]
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; [illegible]49[illegible]; AA[illegible];
DR   InterPro; IPR00[illegible]99;
DR   InterPro; IPR00[illegible];
DR   Pfam; PF01852; START; 1.
DR   PRINTS; PR00[illegible]78; STARPROTEIN.
KW   Lipid-binding; Lipid transport; Transport; [illegible]
KW   Mitochondrion; Transit peptide.
FT   TRANSIT      [illegible]       MITOCHONDRION (POTENTIAL).
FT   CHAIN        [illegible]  284   STEROIDOGENIC ACUTE REGULATORY PROTEIN.
SQ   SEQUENCE   284 AA;  [illegible] MW;  [illegible]

  Query Match             [illegible]
  Best Local Similarity   [illegible]
  Matches [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible];

Qy          1 ALI[illegible]LNN[illegible]GW[illegible] 12
              [illegible]
Db         85 ALI[illegible]LSNQ[illegible]GW[illegible] 96


RESULT 6
STAR_BOVIN
ID   STAR_BOVIN     STANDARD;      [illegible]   284 AA.
AC   Q28918;
DT   01-NOV-1997 (Rel. 35, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN PRECURSOR [illegible]
GN   STAR.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; [illegible]
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; [illegible]
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Corpus luteum;
RX   MEDLINE=96011627; PubMed=[illegible];
RA   Hartung S., Rust W., [illegible] M., [illegible]
RT   "Molecular cloning and in vivo expression of the bovine [illegible]
RT   acute regulatory protein."
RL   Biochem. Biophys. Res. Commun. [illegible]
RN   [2]
RP   REVISIONS TO [illegible]
RA   Hartung S., Rust W., [illegible] M., [illegible]
RL   Submitted (MAY-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: PLAYS A KEY ROLE IN STEROID HORMONE SYNTHESIS BY
CC       ENHANCING THE METABOLISM OF [illegible]
CC       TRANSPORTER THAT BINDS TO AND TRANSPORT [illegible]
CC       INTERMEMBRANE SPACE OF THE MITOCHONDRION (PROBABLE).
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL.
CC   -!- TISSUE SPECIFICITY: CORPUS LUTEUM AND ADRENAL [illegible]
CC   -!- SIMILARITY: BELONGS TO THE STAR FAMILY.
CC   --------------------------------------------------------------------------
```

```
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; [illegible]; AA[illegible]
DR   EMBL; [illegible]; AA[illegible]
DR   InterPro; IPR[illegible]
DR   InterPro; IPR[illegible]
DR   Pfam; PF01852; START; 1.
DR   PRINTS; PR[illegible]; STARPROTEIN.
KW   Lipid-binding; Lipid transport; Transport; Steroidogenesis;
KW   Mitochondrion; Transit peptide.
FT   TRANSIT      [illegible]       MITOCHONDRION (POTENTIAL).
FT   CHAIN        [illegible]  28[illegible]   STEROIDOGENIC ACUTE REGULATORY PROTEIN.
FT   CONFLICT     [illegible]       [illegible] (IN REF. [illegible]; NO M[illegible] SEQUENCE).
FT   CONFLICT     [illegible]       [illegible] K (IN REF. [illegible]; NO M[illegible] SEQUENCE).
SQ   SEQUENCE   [illegible] AA;  [illegible] MW;  [illegible]

  Query Match             [illegible]; Score [illegible]; Length 28[illegible];
  Best Local Similarity   [illegible]; Pred. No. [illegible];
  Matches [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible];

Qy          1 A[illegible]NNQ[illegible]GW[illegible]
              [illegible]
Db         [illegible] A[illegible]LSNQ[illegible]GW[illegible]


RESULT 7
STAR_[illegible]
ID   STAR_[illegible]     STANDARD;      [illegible]
AC   [illegible];
DT   [illegible]-NOV-[illegible] (Rel. [illegible], Created)
DT   [illegible]-NOV-[illegible] (Rel. [illegible], Last sequence update)
DT   [illegible]-MAY-[illegible] (Rel. [illegible], Last annotation update)
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN PRECURSOR [illegible]
GN   STAR.
OS   [illegible]
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; [illegible]
OX   NCBI_TaxID=[illegible];
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=[illegible];
RX   MEDLINE=[illegible]; PubMed=[illegible];
RA   [illegible] N., [illegible]
RA   [illegible]
RT   "[illegible] and bovine steroidogenic acute regulatory protein (StAR)
RT   [illegible]";
RL   Endocrinology [illegible]
RN   [2]
RP   SEQUENCE [illegible] FROM N.A.
RC   TISSUE=[illegible];
RX   MEDLINE=[illegible]; PubMed=[illegible];
RA   [illegible] K., [illegible] H.A., [illegible] M.,
RA   [illegible]
RT   "Regulation of [illegible] cell steroidogenic acute regulatory
RT   protein (StAR) [illegible] with [illegible] synergism with
RT   [illegible] hormone or protein kinase A agonists.";
RL   Endocrinology [illegible]
CC   -!- FUNCTION: PLAYS A KEY ROLE IN STEROID HORMONE SYNTHESIS BY
CC       ENHANCING THE METABOLISM OF CHOLESTEROL INTO PREGNENOLONE.
CC       TRANSPORTER THAT BINDS TO AND TRANSPORT [illegible] THROUGH THE
CC       INTERMEMBRANE SPACE OF THE MITOCHONDRION (PROBABLE).
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN ADULT AND FETAL STEROIDOGENIC
CC       TISSUES [illegible] ADULT TESTES AND OVARIES AND ADULT ADRENAL
CC       [illegible] AS WELL AS [illegible] TISSUES OF PREGNANCY, INCLUDING
CC       DEVELOPING [illegible] CORPUS LUTEUM AND PREGNANCY [illegible]
CC       THE FETAL OVARY.
CC   -!- SIMILARITY: BELONGS TO THE STAR FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; U53020; AAB04953.1; -.
DR   EMBL; U72195; AAB30555.1; -.
DR   InterPro; IPR000799; -.
DR   InterPro; IPR002913; -.
DR   Pfam; PF01852; START; 1.
DR   PRINTS; PR00978; STARPROTEIN.
KW   Lipid-binding; Lipid transport; Transport; Steroidogenesis;
KW   Mitochondrion; Transit peptide.
FT   TRANSIT       1      ?       MITOCHONDRION (POTENTIAL).
FT   CHAIN         ?    285       STEROIDOGENIC ACUTE REGULATORY PROTEIN.
FT   CONFLICT    188    188       G -> R (IN REF. 2).
     SEQUENCE   285 AA;  31833 MW;  67DCA9 A[illegible]21519B8 CRC64;

  Query Match             81.8%;  Score 54;  DB 1;  Length 285;
  Best Local Similarity   83.3%;  Pred. No. [illegible]015;
  Matches   10;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 ALGILNNQEGWK 12
              || ||:||||||
Db         86 ALDILSNQEGWK 97


RESULT   8
UBC3_SCHPO
ID   UBC3_SCHPO     STANDARD;      PRT;   15[illegible]  AA.
AC   P40984;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   UBIQUITIN-CONJUGATING ENZYME E2-18 KDA (EC 6.3.2.19)
DE   (UBIQUITIN-PROTEIN LIGASE HUS5) (UBIQUITIN CARRIER PROTEIN HUS5).
GN   HUS5 OR SPAC30D11.13.
OS   Schizosaccharomyces pombe (Fission yeast).
OC   Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC   Schizosaccharomycetales; Schizosaccharomycetaceae;
OC   Schizosaccharomyces.
OX   NCBI_TaxID=4896;
RN   [1]
     SEQUENCE FROM N.A.
     MEDLINE=95286695; PubMed=7768995
     Al-Khodairy F., Enoch T., Hagman I.M., Carr A.M.;
RT   "The Schizosaccharomyces pombe hus5 gene encodes a ubiquitin
RT   conjugating enzyme required for normal mitosis.";
RL   J. Cell Sci. 108:475-486(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=972;
RA   Pearson D., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;
RL   Submitted (NOV-1995) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO
CC       OTHER PROTEINS. REQUIRED FOR EFFICIENT RECOVERY FROM DNA DAMAGE OR
CC       S-PHASE ARREST AND NORMAL MITOSIS. THIS MAY BE AS PART OF A
CC       CHECKPOINT INDEPENDENT RECOVERY PROCESS.
CC   -!- CATALYTIC ACTIVITY: ATP + UBIQUITIN + PROTEIN LYSINE = AMP +
CC       PYROPHOSPHATE + PROTEIN N-UBIQUITYLLYSINE.
CC   -!- PATHWAY: SECOND STEP IN UBIQUITIN CONJUGATION.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR.
CC   -!- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR
CC       UBIQUITIN-THIOLESTER FORMATION.
CC   -!- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
CC       STRONGEST, TO YEAST UBC9.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; X81846; CAA57438.1; -.
DR   EMBL; Z67961; CAA91899.1; -.
DR   HSSP; P50550; 1A3S.
DR   InterPro; IPR000608; -.
DR   Pfam; PF00179; UQ_con; 1.
DR   PROSITE; PS00183; UBIQUITIN_CONJUGAT_1; 1.
DR   PROSITE; PS50127; UBIQUITIN_CONJUGAT_2; 1.
KW   Ubiquitin conjugation; Ligase; DNA repair; Nuclear protein;
KW   Multigene family.
FT   BINDING      93     93       UBIQUITIN (BY SIMILARITY).
SQ   SEQUENCE   1[illegible]7 AA;  17989 MW;  4BF3A63892 3D27F CRC64;

  Query Match             71.2%;  Score 47;  DB 1;  Length 157;
  Best Local Similarity   72.7%;  Pred. No. [illegible];
  Matches    8;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 LSILNNQEGWK 12
              | ||| : |:
Db         94 LSILNEEECWK 104


RESULT   9
GUN_BACS1
ID   GUN_BACS1      STANDARD;      PRT;   8[illegible]  AA.
AC   P06564;
DT   01-JAN-1988 (Rel. 06, Created)
DT   01-JAN-1988 (Rel. 06, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   ENDOGLUCANASE PRECURSOR (EC 3.2.1.4) (ENDO-1,4-BETA-GLUCANASE)
DE   (ALKALINE CELLULASE).
OS   Bacillus sp. (strain 1139).
OC   Bacteria; Firmicutes; Bacillus/Clostridium group;
OC   Bacillus/Staphylococcus group; Bacillus.
OX   NCBI_TaxID=1411;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87085442; PubMed=3098909;
RA   Fukumori F., Kudo T., Narahashi Y., Horikoshi K.;
RT   "Molecular cloning and nucleotide sequence of the alkaline cellulase
RT   gene from the alkalophilic Bacillus sp. strain 1139.";
RL   J. Gen. Microbiol. 132:2329-2335(1986).
CC   -!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF 1,4-BETA-D-GLUCOSIDIC
CC       LINKAGES IN CELLULOSE.
CC   -!- MISCELLANEOUS: ALKALOPHILIC BACILLUS SP., STRAIN 1139, PRODUCES
CC       ONE ALKALINE CELLULASE (PH OPTIMUM [illegible]). THIS STRAIN IS NOT A TRUE
CC       CELLULOLYTIC MICRO-ORGANISM BECAUSE THE ENZYME IS UNABLE TO
CC       HYDROLYSE NATIVE CELLULOSE.
CC   -!- SIMILARITY: BELONGS TO CELLULASE FAMILY A (FAMILY 5 OF GLYCOSYL
CC       HYDROLASES).
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; M15743; AAA22305.1; -.
DR   EMBL; D00066; BAA00045.1; -.
DR   PIR; A29003; A29003.
DR   InterPro; IPR001547; -.
DR   Pfam; PF00150; cellulase; 1.
DR   PROSITE; PS00[illegible] [illegible]YCOSYL HYD[illegible]
KW   Cellulose degradation; Hydrolase; Glycosidase; [illegible]
FT   SIGNAL       [illegible]                     [illegible]TENTIAL
FT   CHAIN        [illegible]                     ENDOGLUCANASE
FT   ACT SITE     [illegible]                     [illegible]N [illegible] (BY SIMILARITY).
FT   ACT SITE     [illegible]                     N[illegible] (BY SIMILARITY).
SQ   SEQUENCE   [illegible] AA;  [illegible] MW;  [illegible]CA4[illegible]DA[illegible]

  Query Match             [illegible]
  Best Local Similarity   [illegible]
  Matches    7;  Conservative  [illegible]  Mismatches  [illegible]

Qy          3 [illegible]TENNQE[illegible]W [illegible]
              [illegible]
Db        200 [illegible]EPNNEE[illegible]W [illegible]8

[illegible]T 10
GUNA_CALSA
ID   GUNA_CALSA     STANDARD;      PRT;  [illegible] AA.
AC   P22534;
DT   01-AUG-1991 (Rel. 19, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   15-JUL-1998 (Rel. 3[illegible], Last annotation update)
DE   ENDOGLUCANASE A PRECURSOR (E[illegible]) (ENDO[illegible]ANA[illegible] A)
DE   (CELLULASE A).
GN   CELA.
OS   Caldocellum saccharolyticum (Caldicellulosiruptor [illegible]
OC   Bacteria; Firmicutes; Bacillus/Clostridium group;
OC   Thermoanaerobacter group; Caldicellulosiruptor.
OX   NCBI_TaxID=44001;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=[illegible]; PubMed=[illegible];
RA   Te'o V.S., Saul D.J., Bergquist P.L.;
RT   "celA, another gene coding for a multidomain [illegible]
RT   extreme thermophile Caldocellum saccharolyticum.";
RL   Appl. Microbiol. Biotechnol. [illegible]
RN   [2]
RP   SEQUENCE OF [illegible]74, FROM N.A.
RX   MEDLINE=9124[illegible]; PubMed=[illegible];
RA   Leuthi E., Jasmat N.B., Grayling R.A., [illegible]
RT   "Cloning, sequence analysis, and expression [illegible]
RT   gene coding for a beta-mannanase from the [illegible]
[illegible]  bacterium 'Caldocellum saccharolyticum'."
[illegible]  Appl. Environ. Microbiol. [illegible]
CC   -!- FUNCTION: THE N-TERMINAL DOMAIN OF CELA [illegible]
CC       ENDOGLUCANASE ACTIVITY ON CARBOXYMETHYLCELLULOSE [illegible]
CC       DOMAIN PROBABLY ACT SYNERGISTICALLY [illegible]
CC       CELLULOSE.
CC   -!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS [illegible]
CC       LINKAGES IN CELLULOSE.
CC   -!- PTM: THE LINKER REGION (ALSO TERMED "HINGE") MAY BE [illegible]
CC       SITE FOR PROTEOLYSIS.
CC   -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO [illegible]
CC       E (FAMILY [illegible] GLYCOSYL HYDROLASES).
CC   -!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO [illegible] FAMILY
CC       L (FAMILY 48 OF GLYCOSYL HYDROLASES).
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. [illegible]
CC   between the Swiss Institute of Bioinformatics [illegible]
CC   the European Bioinformatics Institute. There are [illegible]
CC   use by non-profit [illegible]
CC   modified and this statement is not removed. [illegible]
CC   entities requires a license agreement (See http://[illegible]/
CC   or send an email to [illegible]).
CC   --------------------------------------------------------------------------
DR   EMBL; L32742; AAA[illegible]
DR   EMBL; M36063; AAA[illegible]
DR   EMBL; L01257; -; NOT_ANNOTATED_CDS.
DR   PIR; A43[illegible]45; A43[illegible]
DR   HSSP; P26221; [illegible]
```

```
DR   [illegible]
DR   [illegible]
DR   [illegible]
DR   [illegible]
DR   Pfam; PF[illegible]
DR   Pfam; PF[illegible]
DR   PRINTS; [illegible]
DR   PROSITE; PS00[illegible]; GLYCOSYL_HYDROL_F[illegible]
DR   PROSITE; PS00[illegible]; GLYCOSYL_HYDROL_F[illegible]
KW   Cellulose degradation; Hydrolase; [illegible] Signal; Repeat
FT   SIGNAL       [illegible]                POTENTIAL.
FT   CHAIN        [illegible]                ENDOGLUCANASE A.
FT   DOMAIN       [illegible]                CATALYTIC.
FT   DOMAIN       [illegible]                LINKER ("HINGE") (PRO-THR BOX).
FT   DOMAIN       [illegible]                CELLULOSE-BINDING (BY SIMILARITY).
FT   DOMAIN       [illegible]                LINKER ("HINGE") (PRO-THR BOX).
FT   DOMAIN       [illegible]                CELLULOSE-BINDING (BY SIMILARITY).
FT   DOMAIN       [illegible]                LINKER ("HINGE") (PRO-THR BOX).
FT   DOMAIN       [illegible]                CATALYTIC.
FT   ACT_SITE     [illegible]                BY SIMILARITY.
FT   ACT_SITE     [illegible]                BY SIMILARITY.
FT   ACT_SITE     [illegible]                BY SIMILARITY.
FT   CONFLICT     [illegible]                [illegible] A (IN [illegible]).
SQ   SEQUENCE   [illegible] AA;  [illegible] MW;  [illegible]

  Query Match             [illegible]
  Best Local Similarity   [illegible]
  Matches  [illegible]  Conservative  [illegible]  Mismatches  [illegible]  Indels  [illegible]  Gaps  [illegible]

Qy          [illegible]NNQ[illegible]W [illegible]

Db        [illegible]IVNNQ[illegible]W [illegible]

RES[illegible]
KETA[illegible]
ID   KETA[illegible]      STANDARD;      PRT;   [illegible] AA.
AC   [illegible]
DT   [illegible] (Rel. [illegible], Created)
DT   [illegible] (Rel. [illegible], Last sequence update)
DT   [illegible] (Rel. [illegible], Last annotation update)
DE   RNA [illegible]
GN   KETA [illegible]
OS   Escherichia [illegible]
OC   Bacteria; [illegible] Enterobacteriaceae;
OC   Escherichia.
OX   NCBI_TaxID=[illegible];
RN   [illegible]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K[illegible] MG[illegible];
RX   MEDLINE=[illegible]; PubMed=[illegible];
RA   [illegible] Daniels D.L.;
RA   [illegible]
RT   "Analysis of the Escherichia coli genome [illegible] DNA sequence of the
RT   [illegible]"
RL   Nucleic Acids Res. [illegible]
RN   [illegible]
RP   FUNCTION.
RX   MEDLINE=[illegible]; PubMed=[illegible];
RA   Spinelli [illegible], Kierzek R., Turner D.H., Phizicky E.M.;
RT   "[illegible] phosphate implicated in its
RT   removal from [illegible] tRNA during splicing in yeast.";
RL   J. Biol. Chem. [illegible]
CC   -!- FUNCTION: REMOVES THE [illegible] PHOSPHATE FROM [illegible]RNA VIA AN INTERMEDIATE IN
CC       WHICH THE PHOSPHATE IS ADP-RIBOSYLATED BY NAD FOLLOWED BY A
CC       PRESUMED TRANSESTERIFICATION TO RELEASE THE RNA AND GENERATE ADP-
CC       RIBOSE [illegible] PHOSPHATE (APPR [illegible]). MAY FUNCTION AS AN ADP-
CC       RIBOSYLASE.
CC   -!- SIMILARITY: BELONGS TO THE KPTA / TPT1 FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; U14003; AAA97227.1; ALT_INIT.
DR   EMBL; AE000503; AAC77287.1; ALT_INIT.
DR   EcoGene; EG12570; kptA.
DR   InterPro; IPR002745; -.
DR   Pfam; PF01885; DUF60; 1.
KW   Transferase; NAD.
SQ   SEQUENCE   184 AA;  20529 MW;  2C654B[illegible]2E7494 CRC64;

  Query Match             57.6%;  Score 38;  DB 1;  Length 184;
  Best Local Similarity   54.5%;  Pred. No. 7.4;
  Matches    6;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        1 ALGILNNQEGW 11
            |:||: ::|||
Db       26 AIGIVLDREGW 36


RESULT 12
BGL1_BACSU
ID   BGL1_BACSU     STANDARD;      PRT;   469 AA.
AC   P40740; O32287;
DT   01-FEB-1995 (Rel. 31, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   BETA-GLUCOSIDASE (EC 3.2.1.21) (GENTIOBIASE) (CELLOBIASE) (BETA-D-
DE   GLUCOSIDE GLUCOHYDROLASE) (AMYGDALASE).
GN   BGLH OR N17D.
OS   Bacillus subtilis.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group;
OC   Bacillus/Staphylococcus group; Bacillus.
OX   NCBI_TaxID=1423;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=168 / MARBURG;
RX   MEDLINE=95189730; PubMed=7883710;
RA   le Coq D.G., Lindner C., Krueger S., Steinmetz M., Stuelke J.;
RT   "New beta-glucoside (bgl) genes in Bacillus subtilis: the bglP gene
RT   product has both transport and regulatory functions similar to those
RT   of BglF, its Escherichia coli homolog.";
RL   J. Bacteriol. 177:1527-1535(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=168 / BGSC1A1;
RX   MEDLINE=95219088; PubMed=7704263;
RA   Yoshida K.-I., Sano H., Seki S., Oda M., Fujimura M., Fujita Y.;
RT   "Cloning and sequencing of a 29 kb region of the Bacillus subtilis
RT   genome containing the hut and wapA loci.";
RL   Microbiology 141:337-343(1995).
CC   -!- CATALYTIC ACTIVITY: HYDROLYSIS OF TERMINAL, NON-REDUCING BETA-D-
CC       GLUCOSE RESIDUES WITH RELEASE OF BETA-D-GLUCOSE.
CC   -!- SIMILARITY: BELONGS TO FAMILY 1 OF GLYCOSYL HYDROLASES.
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; Z34526; CAA84287.1; -.
DR   EMBL; D31856; BAA06653.1; -.
DR   EMBL; D29985; BAA06257.1; -.
DR   EMBL; Z99124; CAB15962.1; ALT_INIT.
DR   PIR; S47175; S47175.
DR   HSSP; P31546; 1PBG.
DR   SubtiList; BG10935; bglH.
DR   InterPro; IPR001360; -.
DR   Pfam; PF00232; Glyco_hydro_1; 1
DR   PRINTS; PR00131; GLHYDRLASE1.
DR   PROSITE; PS00572; GLYCOSYL_HYDROL_F1_1; 1.
DR   PROSITE; PS00653; GLYCOSYL_HYDROL_F1_2; 1.
KW   Hydrolase; Glycosidase; Cellulose degradation.
FT   ACT_SITE    175    175       PROTON DONOR (POTENTIAL).
FT   ACT_SITE    368    368       NUCLEOPHILE (BY SIMILARITY).
FT   CONFLICT    402    402       A -> [illegible] (IN REF. 1).
SQ   SEQUENCE   469 AA;  53289 MW;  40DCA25F964[illegible] CRC64;

  Query Match             57.6%;  Score 38;  DB 1;  Length 469;
  Best Local Similarity   54.5%;  Pred. No. [illegible];
  Matches    6;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy        2 LGILNNQEGWK 12
            ||:: |  : |
Db      134 LGLVKNYGGWK 144


RESULT 13
TBE_HUMAN
ID   TBE_HUMAN      STANDARD;      PRT;   475 AA.
AC   Q9UJT0;
DT   01-OCT-2000 (Rel. 40, Created)
DT   01-OCT-2000 (Rel. 40, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   TUBULIN EPSILON CHAIN (EPSILON TUBULIN).
GN   TUBE1 OR TUBE.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20089049; PubMed=10620804;
RA   Chang P., Stearns T.;
RT   "Delta-tubulin and epsilon-tubulin: two new human centrosomal tubulins
RT   reveal new aspects of centrosome structure and function.";
RL   Nat. Cell Biol. 2:30-35(2000).
CC   -!- SUBCELLULAR LOCATION: CENTROSOMAL; ASSOCIATED WITH PERICENTRIOLAR
CC       MATERIAL.
CC   -!- SIMILARITY: BELONGS TO THE TUBULIN FAMILY.
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; AF201334; AAF09585.1; -.
KW   Microtubules; GTP-binding.
FT   NP_BIND     148    154       GTP (POTENTIAL).
SQ   SEQUENCE   475 AA;  52931 MW;  [illegible]BAEAF780 CRC64;

  Query Match             57.6%;  Score 38;  DB 1;  Length 475;
  Best Local Similarity   100.0%;  Pred. No. 20;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        7 NQEGWK 12
            ||||||
Db      376 NQEGWK 381


RESULT 14
YG18_YEAST
ID   YG18_YEAST     STANDARD;      PRT;   6[illegible] AA.
AC   P53207;
```

```
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   HYPOTHETICAL 7[illegible].4 KDA PROTEIN IN [illegible]
GN   YGR013W.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288C;
RX   MEDLINE=974[illegible]; PubMed=[illegible];
RA   Rieger M., Brueckner M., Schaefer M., Mueller-Auer S.;
RT   "Sequence analysis of [illegible] kb [illegible]
RT   chromosome VII.";
RL   Yeast 13:1077-1090(1997).
CC   -----
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----
DR   EMBL; Z72798; CAA969[illegible]; -.
DR   SGD; S0003245; [illegible]
KW   Hypothetical protein.
SQ   SEQUENCE   6[illegible] AA;  [illegible] MW;  [illegible] CRC64;

  Query Match             [illegible]%;  Score [illegible];  DB 1;  Length [illegible];
  Best Local Similarity   [illegible]%;  Pred. No. [illegible];
  Matches  [illegible];  Conservative  [illegible];  Mismatches  [illegible];  Indels  [illegible];  Gaps  [illegible];

Qy          5 LNNQEGWK 12
              |   ||||
Db         13 LSSQEGWK 20


RESULT 15
GUN3_BACS4
ID   GUN3_BACS4     STANDARD;      PRT;   [illegible] AA.
AC   P19570;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   ENDOGLUCANASE 3 PRECURSOR (EC 3.2.1.4) (ENDO-1,4-BETA-GLUCANASE)
DE   (CELLULASE).
GN   [illegible]EL[illegible].
OS   Bacillus sp. (strain N-4 / JCM [illegible]).
OC   Bacteria; Firmicutes; Bacillus/Clostridium group;
OC   Bacillus/Staphylococcus group; Bacillus.
OX   NCBI_TaxID=141[illegible];
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=893[illegible]; PubMed=[illegible];
RA   Fukumori F., Kudo T., Sashihara N., Nagata Y., [illegible]
RT   "The third cellulase of alkalophilic Bacillus sp. [illegible]
RT   evolutionary relationships within the [illegible]";
RL   Gene 76:289-298(1989).
CC   -!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF 1,4-BETA-D-GLUCOSIDIC
CC       LINKAGES IN CELLULOSE.
CC   -!- SIMILARITY: BELONGS TO THE CELLULASE FAMILY A (FAMILY 5 OF GLYCOSYL
CC       HYDROLASES).
CC   -----
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----
DR   EMBL; M[illegible]; AAA[illegible]; -.
DR   PIR; [illegible]
DR   [illegible]
DR   PFAM; PF00150; cellulase; 1.
DR   PROSITE; [illegible]
KW   Cellulose degradation; Hydrolase; Glycosidase; Signal.
FT   SIGNAL       [illegible]
FT   CHAIN        [illegible]       ENDOGLUCANASE 3.
FT   ACT_SITE     [illegible]       PROTON DONOR (BY SIMILARITY).
FT   ACT_SITE     [illegible]       NUCLEOPHILE (BY SIMILARITY).
SQ   SEQUENCE   [illegible] AA;  [illegible] MW;  [illegible] CRC64;

  Query Match             [illegible]%;  Score [illegible];  DB 1;  Length [illegible];
  Best Local Similarity   [illegible]%;  Pred. No. [illegible];
  Matches  [illegible];  Conservative  [illegible];  Mismatches  [illegible];  Indels  [illegible];  Gaps  [illegible];

Qy          5 [illegible]NNQEGWK [illegible]

Db        [illegible]NQEGWK [illegible]


Search completed: Aug [illegible], 2001
Job time : [illegible]
```

```
                    GenCore version 4.5
          Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on:          August 22, 2001, 14:24:06 ; Search time 55.[illegible] Seconds
                                             (without alignments)
                                             102.784 Million cell updates/sec

Title:           US-08-659-254-2
Perfect score:   1429
Sequence:        1 MPLATFKLCAGSFYRHMRNM...........[illegible] 276

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

[illegible]ched:        197339 seqs, 20590346 residues

Total number of hits satisfying chosen parameters:      197339

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       Issued_Patents_AA:*
                1:  /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
                2:  /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
                3:  /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
                4:  /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
                5:  /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
                6:  /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                              SUMMARIES
                %
Result          Query
   No.  Score   Match Length DB  ID                                Description
------------------------------------------------------------------------------------
    1    1429   100.0    276  4  US-08-659-254-2                   Sequence 2, Appli
    2    1429   100.0    284  2  US-08-538-960-2                   Sequence 2, Appli
    3  1258.5    88.1    285  4  US-08-659-254-18                  Sequence 18, Appl
    4  1252.5    87.6    285  1  US-08-410-[illegible]-2           Sequence 2, Appli
    5     374    26.2    445  2  US-08-69[illegible]-8-4B-6        Sequence 6, Appli
    6     103     7.2     19  2  US-08-538-960-5                   Sequence 5, Appli
    7     103     7.2     19  4  US-08-659-254-5                   Sequence 5, Appli
    8    99.5     7.0    291  4  US-09-328-[illegible]9-2          Sequence 2, Appli
    9    96.5     6.8   2763  3  US-08-49[illegible]-14-2          Sequence 2, Appli
   10      92     6.4    527  4  US-0[illegible]-98[illegible]-18-9  Sequence [illegible], Appli
   11    87.5     6.1    255  1  US-08-45[illegible]-4-4           Sequence 4, Appli
   12    87.5     6.1    255  1  US-08-45[illegible]-3-4           Sequence 4, Appli
   13      83     5.8    553  1  US-08-[illegible]-1-4             Sequence 4, Appli
   14      80     5.6     14  2  US-08-538-960-3                   Sequence 3, Appli
   15      80     5.6     14  4  US-08-659-254-3                   Sequence 3, Appli
   16      80     5.6   2890  4  US-09-41[illegible]-4-67          Sequence 67, Appli
   17      80     5.6   3798  3  US-09-33[illegible]-19-6          Sequence 6, Appli
   18    79.5     5.6    222  1  US-08-79[illegible]-4B-2          Sequence 2, Appli
   19    79.5     5.6    222  4  US-09-39[illegible]-21-6          Sequence 6, Appli
   20    78.5     5.5    572  2  US-08-74[illegible]-34-3          Sequence 3, Appli
   21    78.5     5.5    572  4  US-09-14[illegible]-9-6           Sequence 6, Appli
   22      77     5.4    213  3  US-09-20[illegible]-8-5           Sequence 5, Appli
   23      77     5.4    244  3  US-09-13[illegible]-82-4          Sequence 4, Appli
   24      77     5.4    434  2  US-08-46[illegible]-12-6          Sequence 6, Appli
   25    76.5     5.4    573  2  US-08-74[illegible]-34-4          Sequence 4, Appli
   26    76.5     5.4    573  4  US-09-14[illegible]-9-7           Sequence 7, Appli
   27    75.5     5.3    765  4  US-08-44[illegible]-818-70        Sequence 70, Appl
   28      75     5.2    600  2  US-[illegible]                    Sequence 4, Appli
   29      75     5.2    600  3  US-[illegible]-4-[illegible]      Sequence 4, Appli
   30      75     5.2    614  1  US-[illegible]-9[illegible]A-2    Sequence 2, Appli
   31      75     5.2    614  2  US-08-[illegible]-2               Sequence 2, Appli
   32      75     5.2    614  3  US-0[illegible]-14[illegible]     Sequence 19, Appli
   33      75     5.2    614  3  US-0[illegible]-14[illegible]     Sequence 2[illegible], Appli
   34      75     5.2    614  3  US-08-14[illegible]               Sequence 2[illegible], Appli
   35      75     5.2    614  3  US-0[illegible]-14[illegible]     Sequence 25, Appli
   36      75     5.2    614  3  US-08-[illegible]14-[illegible]   Sequence 2, Appli
   37      75     5.2    614  5  PCT-US[illegible]-6-2             Sequence 2, Appli
   38      75     5.2    617  2  US-08-[illegible]                 Sequence 6, Appli
   39      75     5.2    617  3  US-09-[illegible]14-[illegible]   Sequence 6, Appli
   40    74.5     5.2    676  4  US-0[illegible]-A-2               Sequence 2, Appli
   41    74.5     5.2   2077  6  522412[illegible]                 Patent No. [illegible]22[illegible]423
   42      74     5.2    614  3  US-0[illegible]-14[illegible]     Sequence 2[illegible], Appli
   43      74     5.2   1832  3  US-09-[illegible]-4-[illegible]   Sequence 4, Appli
   44    73.5     5.1    201  3  US-0[illegible]-72[illegible]     Sequence 2, Appli
   45      73     5.1    942  3  US-0[illegible]-5-[illegible]     Sequence [illegible], Appli

                              ALIGNMENTS

RESULT 1
US-08-659-254-2
; Sequence 2, Application US/08659254
; Patent No. 61945[illegible]
;  GENERAL INFORMATION:
;    APPLICANT:  Stocco, Douglas M.
;    APPLICANT:  Clark, Dr. Barbara J.
;    TITLE OF INVENTION: [illegible]
;    TITLE OF INVENTION: [illegible]
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Akin, Gump, Strauss, Hauer & Feld, L.L.P.
;      STREET:  [illegible] Park [illegible]
;      CITY:  Austin
;      STATE:  TX
;      COUNTRY:  U.S.A.
;      ZIP:  78701
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.[illegible], Version #1.[illegible]
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/659,254
;      FILING DATE:  07-JUN-1996
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/5[illegible]
;      FILING DATE:  04-OCT-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Mayfield, Denise L.
;      REGISTRATION NUMBER:  33,732
;      REFERENCE/DOCKET NUMBER:  43375.[illegible]6
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  512/499-6200
;      TELEFAX:  512/499-6290
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  276 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-659-254-2

  Query Match            100.0%;  Score 1429;  DB 4;  Length 276;
  Best Local Similarity  100.0%;  Pred. No. [illegible];
  Matches  276;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          1 MELATEK[illegible]YE[illegible]MENMK[illegible]PEQAVLAT[illegible]NWEA[illegible]
Db          1 MELATEK[illegible]YE[illegible]MENMK[illegible]EQAV[illegible]A[illegible]NWEA[illegible]

Qy         61 [illegible]ATLY[illegible]SYE[illegible]EVAMQKA[illegible]NNQ[illegible]WEK[illegible]
Db         61 [illegible]EATLY[illegible]SYE[illegible]EVAMQKA[illegible]NNQ[illegible]WEKE[illegible]

Qy        121 [illegible]VVV[illegible]EM[illegible]YEE[illegible]MEAM[illegible]WN[illegible]NVK[illegible]
Db        121 [illegible]VVV[illegible]EM[illegible]YEE[illegible]MEAM[illegible]WN[illegible]NVK[illegible]

Qy        181 EDEVSVE[illegible]TE[illegible]VLA[illegible]MATHE[illegible]
Db        181 EDEVSVE[illegible]TKEE[illegible]VLA[illegible]MATHE[illegible]

Qy        241 [illegible]DEK[illegible]WE[illegible]TNQV[illegible]ANE[illegible]
Db        241 [illegible]DEK[illegible]WE[illegible]TNQV[illegible]ANE[illegible]


[illegible] 2
[illegible] 5 08 [illegible] 2
[illegible]quence [illegible] App[illegible]
  Patent No. 587[illegible]
    GENERAL INFORMATION:
      APPLICANT: [illegible] M
      APPLICANT: Clark, Barbara [illegible]
      TITLE OF INVENTION: [illegible] ME[illegible] AND METH[illegible]
      TITLE OF INVENTION: [illegible]ULATION [illegible]
      NUMBER OF SEQUENCES: 14
      CORRESPONDENCE ADDRESS:
        ADDRESSEE: Ar[illegible]
        STREET: 1000 Frost Bank Plaza [illegible]
        CITY: Austin
        STATE: TX
        COUNTRY: U.S.A.
        ZIP: 787[illegible]
      COMPUTER READABLE FORM:
        MEDIUM TYPE: Floppy disk
        COMPUTER: IBM [illegible]
        OPERATING SYSTEM: [illegible]/MS[illegible]
        SOFTWARE: PatentIn [illegible] #[illegible] Ver[illegible] #[illegible]
      CURRENT APPLICATION DATA:
        APPLICATION NUMBER: [illegible]/08/[illegible]
        FILING DATE: [illegible]
        CLASSIFICATION: [illegible]
      ATTORNEY/AGENT INFORMATION:
        NAME: Mayfield, De[illegible]
        REGISTRATION NUMBER: [illegible]
        REFERENCE/DOCKET NUMBER: 4[illegible]M
      TELECOMMUNICATION INFORMATION:
        TELEPHONE: 512/[illegible]
        TELEFAX: [illegible]
    INFORMATION FOR SEQ ID NO: [illegible]
      SEQUENCE CHARACTERISTICS:
        LENGTH: 284 [illegible]
        TYPE: amino [illegible]
        STRANDEDNESS: [illegible]
        TOPOLOGY: [illegible]
      MOLECULE TYPE: [illegible]
[illegible]

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches [illegible] Conservative [illegible] Mismatches [illegible]

Qy          1 MELATEKQ[illegible]YEEMENMK[illegible]EQAVLAT[illegible]NWEA[illegible]
Db          1 MELATEKQ[illegible]YEEMENMK[illegible]EQAVLAT[illegible]NWEA[illegible]

Qy         61 QEEATLY[illegible]EVAMQKA[illegible]NNQ[illegible]WEK[illegible]
```

```
[illegible] VAMQKA[illegible]NNQ[illegible]WEKE[illegible]

Qy [illegible] VVV[illegible]EM[illegible]MEAM[illegible]WNENVKE[illegible]AAAA[illegible]

Db [illegible] VVV[illegible]EM[illegible]MEAM[illegible]WNENVKE[illegible]AAAA[illegible]

Qy [illegible]MATHE[illegible]

Db [illegible]MATHE[illegible]

Qy [illegible]WE[illegible]TNQV[illegible]ANE[illegible]

Db [illegible]WE[illegible]TNQV[illegible]ANE[illegible]


[illegible]
    GENERAL INFORMATION:
      APPLICANT: [illegible] M
      APPLICANT: [illegible]
      TITLE OF INVENTION: [illegible] AND [illegible] REGULATION
      TITLE OF INVENTION: [illegible]
      NUMBER OF SEQUENCES: [illegible]
      CORRESPONDENCE ADDRESS:
        ADDRESSEE: Ar[illegible]
        STREET: [illegible] Austin
        CITY: Austin
        STATE: TX
        COUNTRY: U.S.A.
        ZIP: [illegible]
      COMPUTER READABLE FORM:
        MEDIUM TYPE: [illegible]
        COMPUTER: IBM [illegible] compatible
        OPERATING SYSTEM: [illegible]
        SOFTWARE: PatentIn Release #1.0, Version #[illegible]
      CURRENT APPLICATION DATA:
        APPLICATION NUMBER: [illegible]
        FILING DATE: [illegible]
        CLASSIFICATION: [illegible]
      PRIOR APPLICATION DATA:
        APPLICATION NUMBER: [illegible]
        FILING DATE: [illegible]
      ATTORNEY/AGENT INFORMATION:
        NAME: Mayfield, Denise L
        REGISTRATION NUMBER: [illegible]
        REFERENCE/DOCKET NUMBER: [illegible]
      TELECOMMUNICATION INFORMATION:
        TELEPHONE: [illegible]
        TELEFAX: [illegible]
    INFORMATION FOR SEQ ID NO: [illegible]
      SEQUENCE CHARACTERISTICS:
        LENGTH: [illegible]
        TYPE: [illegible]
        TOPOLOGY: [illegible]
[illegible]

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches [illegible] Mismatches [illegible] Indels [illegible] Gaps [illegible]

Qy [illegible] MEA[illegible]MENMK[illegible]AVLA[illegible]NWEA[illegible]
Db [illegible] MEA[illegible]MENMK[illegible]AVMA[illegible]NEEA[illegible]

Qy [illegible]VAMQKA[illegible]NNQ[illegible]WEK[illegible]

Db [illegible]AMQKA[illegible]WEK[illegible]KVME[illegible]

[illegible]MEAM[illegible]WNENVKE[illegible]AAAA[illegible]
```

```
              [illegible]
Db      121 PLFVVVDQPMPRLYFELVERMFAMCFWNPNVFL [illegible]

Qy      180 PRDFVSVRCTEKRRG [illegible] VLAGMATHF [illegible]
              [illegible]
Db      181 FRDFVSV [illegible] AKR [illegible] VLAGMAT [illegible]

Qy      240 WLLSIDLKGWLPKTIINQVLSQTQIEFANHLRK [illegible] 276
              [illegible]
Db      241 WLLSIDLKGWLPKSIINQVLSQTQVDFANHLRKRL [illegible] S 277


RESULT    4
US-~~08-410-540-2~~
; Sequence 2, Application US/08410540
; Patent No. 5807678
;  GENERAL INFORMATION:
;    APPLICANT:  Miller, Walter L.
;    APPLICANT:  Lin, Dong
;    APPLICANT:  Strauss III, Jerome F
;    TITLE OF INVENTION:  IDENTIFICATION OF GENE MUTATIONS
;    TITLE OF INVENTION:  ASSOCIATED WITH CONGENITAL LIPOID ADRENAL HYPERPLASIA
;    NUMBER OF SEQUENCES:  30
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Cooley Godward Castro Huddleson & Tatum
;      STREET:  5 Palo Alto Square
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  US
;      ZIP:  94306-2155
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/410,[illegible]
;      FILING DATE:  23-MAR-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Neeley, Richard L.
;      REGISTRATION NUMBER:  30,092
;      REFERENCE/DOCKET NUMBER:  UCAL-228/[illegible]
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415 853 5070
;      TELEFAX:  415 857 0663
;      TELEX:  380816COOLEYPA
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  285 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-410-540-2


  Query Match             87.6%;  Score [illegible];  DB 1;  Length 285;
  Best Local Similarity   87.0%;  Pred. No. [illegible]-135;
  Matches  241;  Conservative   18;  Mismatches   17;  Indels   [illegible];  Gaps   [illegible];

Qy        1 MFLATFKLCAGSSYRHMRNMKGLRHQAVLAI [illegible] LNWRAL [illegible] SSPG-WM [illegible] VKRRSSLLG 59
              [illegible]
Db        1 MLLATFKLCAGSSYRHMRNMKGLRQQAVMAI [illegible] LNFRAL [illegible] WINQVRRSSLLG [illegible]

Qy       60 SQ [illegible] FATLYSDQFLSY [illegible] FVAMQFAL [illegible] VKE 119
              [illegible]
Db       61 SRLEETLYSDQFLAYLQQCFFAMQFAL [illegible] VKE [illegible]

Qy      120 RLFVVVDQPMPRLYEELVERMFAMCEWN [illegible]
              [illegible]
Db      121 RLFVVVDQPM [illegible] YEELVERMEAMCEWN [illegible]

Qy      180 PRDFVSVRCTKRRGSTCVLAGMATHFGEM [illegible] VIRAEHGPT [illegible] MVLHPLAGSLSKTKL [illegible]
```

```
              [illegible]
Db      181 PRDFVSVR [illegible] AKRRGS [illegible] VLAGMD [illegible] SKIKLL 240

Qy      240 WLLSI [illegible] KIIINQVL [illegible] ANHLRKR [illegible] 276
              [illegible]
Db      241 WLLS [illegible] TINQVL [illegible] ANHLRKRLES 277


RESULT    5
US-08-691-814B-6
; Sequence 6, Application US/08691814B
; Patent No. 598[illegible]
;  GENERAL INFORMATION:
;    APPLICANT:  Rio, Marie Christine
;    APPLICANT:  Tomasetto, Catherine
;    APPLICANT:  Basset, Paul
;    APPLICANT:  Byrne, Jennifer
;    TITLE OF INVENTION:  Isolated Nucleic Acid Molecules Useful
;    TITLE OF INVENTION:  as Leukemia Markers and in Breast Cancer Prognosis
;    NUMBER OF SEQUENCES:  124
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sterne, Kessler, Goldstein & Fox P.L.L.C.
;      STREET:  1100 New York Ave., NW, [illegible]
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20005-3934
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #[illegible]  Version #[illegible]
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/69[illegible]
;      FILING DATE:  31-JUL-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/002,[illegible]
;      FILING DATE:  09-AUG-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Steffe, Eric K.
;      REGISTRATION NUMBER:  36,688
;      REFERENCE/DOCKET NUMBER:  1383.[illegible]
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-371-2600
;      TELEFAX:  202-371-2540
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  445 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-691-814B-6


  Query Match             26.2%;  Score [illegible]74;  DB 2;  Length 445;
  Best Local Similarity   38.5%;  Pred. No. [illegible]e-34;
  Matches   80;  Conservative   42;  Mismatches   86;  Indels   [illegible];  Gaps    0;

Qy       67 YSDQ [illegible] EVAMQKALGILNNQ [illegible] KMV [illegible] KVPRLLVVVE 126
              [illegible]
Db      231 FTAQ [illegible] KEAIAVV [illegible] ILKIFLL 290

Qy      127 QPMDRLYEE [illegible] VDRMEAMGEWNPNVKEIIEV [illegible] AAAAANLV [illegible] VSV 186
              [illegible]
Db      291 CPAELVY [illegible] PEEMVLWNKIVIA [illegible] IAA [illegible] VVNV [illegible]

Qy      187 RCTKRR [illegible] VLAGMATHPGEMPP [illegible] MVLH [illegible] SKTKL [illegible] WLLSI [illegible] 246
              [illegible]
Db      351 RRIERRRDRYLSSCIAISHSAKPPIHK [illegible] IVLK [illegible] EVWILNII [illegible]

Qy      247 KGWLPKI [illegible] NQVLSQTQIEFANHLRKR [illegible]4
              [illegible]
```

```
Db        411 K[illegible]

RESULT 6
US-08-538-[illegible]-5
; Sequence 5, Application US/08538[illegible]
; Patent No. 587[illegible]
;  GENERAL INFORMATION:
;     APPLICANT: [illegible] M.
;     APPLICANT: [illegible]
;     TITLE OF INVENTION: COMPOSITIONS AND METHODS [illegible]
;     TITLE OF INVENTION: [illegible]
;     NUMBER OF SEQUENCES: [illegible]4
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE: [illegible]
;       STREET: [illegible]
;       CITY: Austin
;       STATE: TX
;       COUNTRY: U.S.A.
;       ZIP: 787[illegible]
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE: Floppy disk
;       COMPUTER: IBM PC compatible
;       OPERATING SYSTEM: PC-DOS/MS-DOS
;       SOFTWARE: PatentIn Release #1.0, Version #1.[illegible]
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER: US/08/[illegible]
;       FILING DATE: Concurrently Herewith
;       CLASSIFICATION: [illegible]
;     ATTORNEY/AGENT INFORMATION:
;       NAME: Mayfield, Denise L.
;       REGISTRATION NUMBER: [illegible]
;       REFERENCE/DOCKET NUMBER: [illegible]
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE: [illegible]/49[illegible]
;       TELEFAX: [illegible]
;  INFORMATION FOR SEQ ID NO: [illegible]:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH: [illegible] amino acids
;       TYPE: amino acid
;       STRANDEDNESS: single
;       TOPOLOGY: linear
;     MOLECULE TYPE: peptide
US-08-538-[illegible]-5

  Query Match             [illegible]%;  Score [illegible];  DB [illegible];  Length [illegible];
  Best Local Similarity   [illegible]%;  Pred. No. [illegible];
  Matches [illegible];  Conservative [illegible];  Mismatches [illegible];  Indels [illegible];  Gaps [illegible];

Qy        193 [illegible]VLAGMA[illegible]GEM[illegible] 211
              [illegible]
Db          1 [illegible]VLAGMA[illegible]GEM[illegible] 19


RESULT 7
US-08-659-254-5
; Sequence 5, Application US/0865925[illegible]
; Patent No. 6194[illegible]
;  GENERAL INFORMATION:
;     APPLICANT: [illegible] M.
;     APPLICANT: [illegible]
;     TITLE OF INVENTION: [illegible] AND METHODS [illegible]
;     TITLE OF INVENTION: [illegible]
;     NUMBER OF SEQUENCES: [illegible]
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE: AK[illegible]
;       STREET: 1[illegible]
;       CITY: Austin
;       STATE: TX
;       COUNTRY: U.S.A.
;       ZIP: 78701
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE: Floppy disk
;       COMPUTER: IBM PC compatible
;       OPERATING SYSTEM: PC-DOS/MS-DOS
;       SOFTWARE: [illegible]
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER: [illegible]
;       FILING DATE: [illegible]
;       CLASSIFICATION: [illegible]
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER: [illegible]
;       FILING DATE: [illegible]
;     ATTORNEY/AGENT INFORMATION:
;       NAME: [illegible]
;       REGISTRATION NUMBER: [illegible]
;       REFERENCE/DOCKET NUMBER: [illegible]
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE: [illegible]
;       TELEFAX: [illegible]
;  INFORMATION FOR SEQ ID NO: [illegible]:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH: [illegible] amino acids
;       TYPE: amino acid
;       STRANDEDNESS: single
;       TOPOLOGY: linear
;     MOLECULE TYPE: peptide
US-08-659-254-5

  Query Match             [illegible]%;  Score [illegible];  DB 4;  Length [illegible];
  Best Local Similarity   [illegible]%;  Pred. No. [illegible];
  Matches [illegible];  Conservative [illegible];  Mismatches [illegible];  Indels [illegible];  Gaps [illegible];

Qy        [illegible]VLAGMA[illegible] 211
              [illegible]
Db          1 [illegible]VLAGMA[illegible] 19


RESULT 8
US-[illegible]
; Sequence [illegible]
; Patent No. [illegible]
; GENERAL INFORMATION:
;  APPLICANT: [illegible]
;  APPLICANT: [illegible]
;  APPLICANT: [illegible]
;  TITLE OF INVENTION: [illegible] TRANSFER [illegible]
;  FILE REFERENCE: [illegible]
;  CURRENT APPLICATION NUMBER: US/09/[illegible]
;  CURRENT FILING DATE: [illegible]
;  NUMBER OF SEQ ID NOS: [illegible]
;  SOFTWARE: [illegible]
; SEQ ID NO [illegible]
;  LENGTH: [illegible]
;  TYPE: PRT
;  ORGANISM: [illegible]
;  FEATURE:
;  OTHER INFORMATION: [illegible]
US-[illegible]

  Query Match             [illegible];  Score [illegible];  DB [illegible];  Length [illegible];
  Best Local Similarity   [illegible];  Pred. No. [illegible];
  Matches [illegible];  Conservative [illegible];  Mismatches [illegible];  Indels [illegible];  Gaps [illegible];

Qy        [illegible]AAAAA[illegible]
              [illegible]
Db        [illegible]

Qy        [illegible]VLAGMA[illegible]
              [illegible]
Db        [illegible]

Qy        [illegible]
```

```
              :|:|  :| || |||  ::|:
Db     180 SCVITYLAQVDPKCSLPKWVVNK 202


RESULT   9
US-08-496-944-2
; Sequence 2, Application US/08496944
; Patent No. 6040496
;  GENERAL INFORMATION:
;    APPLICANT:  Law, Marcus D
;    APPLICANT:  Dietz, Jon M
;    TITLE OF INVENTION:  Use of Translationally altered RNA to
;    TITLE OF INVENTION:  Confer Resistance to Maize Dwarf Mosaic Virus and Other
;    TITLE OF INVENTION:  Monocotyledonous Plant Viruses
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  CIBA-Geigy Corporation
       STREET:  7 Skyline Drive
       CITY:  Hawthorne
       STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10532
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30B
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/496,944
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Elmer, James Scott
;      REGISTRATION NUMBER:  36,129
;      REFERENCE/DOCKET NUMBER:  CGC 1814
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2763 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-496-944-2


  Query Match              6.8%;  Score 96.5;  DB 3;  Length 2763;
  Best Local Similarity   20.8%;  Pred. No. 2.23;
  Matches   58;  Conservative   34;  Mismatches   94;  Indels   93;  Gaps   11;

Qy      59 GSQLEATLYSDQELSYIQ-------------------QGEVAMQKALGILNNQEWK 96
           ||  :|| | :  | |:                       |    |   ||:     |
Db    1034 GSTADATKYGNNILVYVASYNPVDSLSQALVELP SVIKVDGRIMKQNIIGIIINITAQK 1093

Qy      97 K    ESQQENGDEVLSKMVPDVGKVIELEVVVQEMDP               LYEELVDEME 141
           |        :  |||        :  |||  | |:'       :|               | ||: |:
Db    1094 KCFVVATNIIENG----VILDIDVGVDEGIKVSADLDVDNRAVLYKRVSISYDGLQEDG 1149

Qy     142 AMGEWNFNV      EEIEVL KIG                   IV         [illegible] 182
            :|    |          | :| || |                   |||| ::
Db    1150 RVGRNKPGTVIRIGKTMKGLQEIPAMIATEAAPMCFAYGLKVITHNVSIHL          - 1201

Qy     183 FVSVRCTKRRGSTC---------VLAGMAIHE[illegible]VIRAEHCP[illegible]MVIHILA IS [illegible]
               :|| ::  |              |:| :                |     ::: :    : ::  | '
Db    1202 ---AKCTVKQARIMMQFELSPFVMAELYKP[illegible]HPQIHEALVKYKLPESVINLPPNA - 1256

Qy     233 PSKTKLTWLLSIDLKGWLPKTIINQVLSQIQH  ANHLR 271
                   |  ::|  ||      |::         :|   :|::
Db    1257 ---------LPRVNLHNWLTARDYNRI   GCS I  EDHVK 1285


RESULT  10
US-08-984-618-9
; Sequence 9, Application US/08984618
; Patent No. 6251647
```

```
;  GENERAL INFORMATION:
;    APPLICANT:  de Lencastre, Herminia
;    APPLICANT:  Tomasz, Alexander
;    TITLE OF INVENTION:  AUXILLIARY GENES AND PROTEINS OF
;    TITLE OF INVENTION:  METHICILLIN RESISTANT BACTERIA AND ANTAGONISTS THEREOF
,    NUMBER OF SEQUENCES:  17
,    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Klauber & Jackson
;      STREET:  411 Hackensack Avenue, 4th Floor
;      CITY:  Hackensack
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07601
,    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/984,618
;      FILING DATE:
;      CLASSIFICATION
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jackson Esq., David A.
;      REGISTRATION NUMBER:  26,742
;      REFERENCE/DOCKET NUMBER:  600-1-[illegible]
,    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201-487-5800
.      TELEFAX:  201-343-1684
;      TELEX:  133521
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  527 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Haemophylus influenzae
US-08-984-618-9


  Query Match              6.4%;  Score [illegible];  DB 4;  Length 527;
  Best Local Similarity   20.9%;  Pred. No. [illegible];
  Matches   62;  Conservative   53;  Mismatches  107;  Indels   74;  Gaps   14;

Qy      11 GSSYRHMKNMKGLRHQAV LATGQELNW[illegible]SKM QVE[illegible]GSLL     GSQLEATL 66
           ||: : :| :||| :   : :|:| :                    :       |:       : : |
Db     196 GSTIQAVEVVKGLNNPELDVAVGDEL                    AQQI[illegible]VQGACNEFEQ A 243

Qy      67 YSDQELSYIQQGEVAMQKALG-    -  INNQEGW- KKESQ[illegible]NGDEVLSKMVPDVGKV 116
            :    ||:  :   |         |||                            ::|       | |          ! |
Db     244 FIKGELIPVEEG          TALGNEGVD[illegible]QWAEK[illegible]SPQA[illegible]VESAEEKESGEV 298

Qy     119 PRLEVVVEQPMDRLYEELVDEMEAM       [illegible]WNPNVEIIKVIE[illegible]GKDIVIHHELA        [illegible]
           |:::     :    |    |             ||:   |                         | [illegible] |:    |
Db     299 FKIQANM[illegible]                [illegible]KVAEM[illegible]VV[illegible]                [illegible]VV[illegible]A[illegible]HA[illegible] 349

Qy     171 ----    AAAAGNLVGPRDF--VSVR[illegible]IKK[illegible]STCVLACMAIHE GEMPEQ              -- 211
                      |    |:::|    :      :  :     |         :           :   :
Db     350 DRAHAEEAYAGDIIGLHNHGTIQIG[illegible]KEI[illegible]IPNEA[illegible]EFRRIRIKDPIKQKQ 409

Qy     212 --SCVIRAIH[illegible]MVIHPLACTPSKIKI[illegible]WLP[illegible]INQVLSQIQIE 265
                 |:::                |     | |    :                :|      .           ::::: :: :
Db     410 LLKGIVQ[illegible]AVQVPEPLLNN[illegible]A[illegible]VI[illegible]              VVGPIKIPYNVI 458


RESULT  11
US-08-459-264-4
; Sequence 4, Application US/08459264
; Patent No. 570534
;  GENERAL INFORMATION:
    APPLICANT: RASMUSSEN, [illegible] A
    APPLICANT: [illegible]
    APPLICANT: [illegible]MAN, Y[illegible]
    TITLE OF INVENTION: [illegible]
    TITLE OF INVENTION: [illegible]
    TITLE OF INVENTION: [illegible]
    NUMBER OF SEQUENCES: [illegible]
    CORRESPONDENCE ADDRESS:
      ADDRESSEE: [illegible]
      STREET: [illegible]
      CITY: New York
      STATE: NY
      COUNTRY: [illegible]
      ZIP: 10[illegible]
    COMPUTER READABLE FORM:
      MEDIUM TYPE: Floppy disk
      COMPUTER: IBM PC compatible
      OPERATING SYSTEM: PC-DOS/MS-DOS
      SOFTWARE: PatentIn Release #1.0, Version #1.[illegible]
    CURRENT APPLICATION DATA:
      APPLICATION NUMBER: US/08/[illegible]
      FILING DATE: [illegible]
      CLASSIFICATION: 4[illegible]
    ATTORNEY/AGENT INFORMATION:
      NAME: Rob[illegible]
      REGISTRATION NUMBER: [illegible]44[illegible]
      REFERENCE/DOCKET NUMBER: [illegible]
    TELECOMMUNICATION INFORMATION:
      TELEPHONE: [illegible]
      TELEFAX: [illegible]
      TELEX: [illegible]
  INFORMATION FOR SEQ ID NO: [illegible]:
    SEQUENCE CHARACTERISTICS:
      LENGTH: [illegible] amino acids
      TYPE: amino acid
      STRANDEDNESS: [illegible]
      TOPOLOGY: [illegible]
    MOLECULE TYPE: peptide
    ORIGINAL SOURCE:
      ORGANISM: [illegible]
    IMMEDIATE SOURCE:
      CLONE: B[illegible]
[illegible]4[illegible]

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches [illegible]

Qy [illegible]
Db [illegible]

Qy      117 V[illegible]EVPREL[illegible]
Db [illegible]

Qy      177 AAAAINLV[illegible]
Db      218 DE[illegible]

Qy      224 M[illegible]VLHPLA[illegible]
Db      16[illegible] VE[illegible]

RESULT 12
US-08-45[illegible]
; Sequence 4, Application US/0[illegible]
; Patent No. 5[illegible]
;  GENERAL INFORMATION:
;    APPLICANT: RASMUSSEN, [illegible] A
;    APPLICANT: [illegible]
```

```
[illegible]
    STATE: NY
[illegible]
    MEDIUM TYPE: Floppy disk
[illegible]
    OPERATING SYSTEM: PC-DOS/MS-DOS
[illegible]
  CURRENT APPLICATION DATA:
[illegible]
  ATTORNEY/AGENT INFORMATION:
[illegible]
  TELECOMMUNICATION INFORMATION:
[illegible]

[illegible]

RESULT [illegible]
[illegible]
; Patent No. [illegible]
;  GENERAL INFORMATION:
;    APPLICANT: HAMM[illegible]
;    APPLICANT: [illegible]
;    APPLICANT: [illegible]
[illegible]
```

```
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  TOWNSEND and TOWNSEND [illegible] and CREW
;      STREET:  Steuart Street Tower, 20th Floor, One Market
;      STREET:  Plaza
;      CITY:  San Francisco
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94105
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  [illegible]
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/106,7[illegible]
;      FILING DATE:  19930813
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Weber, Kenneth A.
;      REGISTRATION NUMBER:  31,677
;      REFERENCE/DOCKET NUMBER:  2307E-14[illegible]
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 543-9600
;      TELEFAX:  (415) 543-5043
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  553 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-106-761-4


  Query Match              5.8%;  Score 8[illegible];  DB 1;  Length 553;
  Best Local Similarity   21.3%;  Pred. No. [illegible]8;
  Matches   56;  Conservative   40;  Mismatches   85;  Indels   82;  Gaps   14;

Qy         37 WPALGDSSPGWMGQVRRPSSLLGSQLEATLYS[illegible]            [illegible]PVAMLK  [illegible]
              |   |  |  |            |:  |     |:  |                |                  :  |    : : : |
Db         63 WVPLMDF[illegible]       YRR[illegible]ANI[illegible]MAAR[illegible]

Qy         85 --ALGILNN--QEGWKKESQQENGDEVL[illegible]KMV[illegible]KVF[illegible]---RLEVVVD[illegible]
                  :  ||      :|  |::|              | :|:  :     |              :|  :  :    ::|
Db        117 FTTCIVTNNWL[illegible]DKRDS---------IAQMM[illegible]

Qy        136 LVDRMEAMGEWNPNV          KEIKVI[illegible]HELAAA  AA[illegible]
              |:|  ::|        ||                :|   :  :      |  :  |    |:|            :  |       |
Db        169 LIDTLKA----KPNFVVFLDDPGSNLE[illegible]IVENIASAL[illegible]A 2[illegible]

Qy        183 EVSVPCTKR---RGSTCVLACMATHPCPMPP[illegible]RAPPCPTCMVLHPLACS[illegible]KIKLI 2[illegible]
                 :  |  |               |      |    |:   | |  |                           |    |           |  |  :  :     :
Db        225 PLPVPCNPNDVCHGYVTVKPGIRLHPVEM                 [illegible]

Qy        240 W-------------LLSIDLKGW 249
              |                    :|:||:||:
Db        271 WRYQIPALASACPRVLAIDMKGY 29[illegible]


RESULT 14
US-08-538-960-3
; Sequence 3, Application US/08538960
; Patent No. 5872230
;  GENERAL INFORMATION:
;    APPLICANT:  Stocco, Douglas M.
;    APPLICANT:  Clark, Barbara J.
;    TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR
;    TITLE OF INVENTION:  REGULATION [illegible]
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Akin, Gump, Strauss, Hauer & Feld, L.L.P.
;      STREET:  1900 Frost Bank Plaza, 81[illegible] Congress Avenue
;      CITY:  Austin
```

```
;      STATE:  TX
;      COUNTRY:  U.S.A.
;      ZIP:  78701
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/[illegible]
;      FILING DATE:  Concurrently Herewith
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Mayfield, Denise L.
;      REGISTRATION NUMBER:  33,732
;      REFERENCE/DOCKET NUMBER:  [illegible]
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  512/499-6200
;      TELEFAX:  512/499-6290
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  14 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-538-960-3


  Query Match              5.6%;  Score [illegible];  DB 2;  Length 14;
  Best Local Similarity   100.0%;  Pred. No. [illegible]5;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        217 AEHCPICMVLHPLA 230
              ||||||||||||||
Db          1 AEHCPICMVLHPLA 14


RESULT 15
US-08-659-254-3
; Sequence 3, Application US/08659254
; Patent No. 6194555
;  GENERAL INFORMATION:
;    APPLICANT:  Stocco, Douglas M.
;    APPLICANT:  Clark, Dr. Barbara J.
;    TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR REGULATION
;    TITLE OF INVENTION:  OF STEROIDOGENESIS
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Akin, Gump, Strauss, Hauer & Feld, L.L.P.
;      STREET:  1900 Frost Bank Plaza, 816 Congress Avenue
;      CITY:  Austin
;      STATE:  TX
;      COUNTRY:  U.S.A.
;      ZIP:  78701
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/659,254
;      FILING DATE:  07-JUN-1996
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/538,960
;      FILING DATE:  04-OCT-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Mayfield, Denise L.
;      REGISTRATION NUMBER:  33,732
;      REFERENCE/DOCKET NUMBER:  [illegible]
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  512/499-6200
```

```
          TELEFAX: [illegible]
   INFORMATION FOR SEQ ID NO: [illegible]
      SEQUENCE CHARACTERISTICS:
          LENGTH: 11 amino acids
          TYPE: amino acid
          STRANDEDNESS: [illegible]
          TOPOLOGY: linear
       MOLECULE TYPE: peptide
[illegible] 08 659 [illegible]

  Query Match             [illegible]
  Best Local Similarity   [illegible]
  Matches   11; [illegible]

Qy        217 AFHGPTCMVL[illegible]A [illegible]
                 |    ||         |
Db          1 AFHGPTCMVL[illegible]A [illegible]


Search completed: [illegible]
Job time: 363 sec
```

```
                    GenCore version [illegible]
           Copyright (c) 1993 - [illegible]

OM protein - protein search, using sw model

Run on:          August 22, 2001, 14:29:08 ; Search time [illegible] Seconds
                                             (without alignments)
                                             7.135 Million cell updates/sec

Title:           US-08-659-254-4
Perfect score:   66
Sequence:        1 ALGILNNQEGWK 12

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.[illegible]

Searched:        412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters:        41[illegible]

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       A_Geneseq_0601:*
               1:   /SIDS8/gcgdata/geneseq/geneseqp/AA[illegible]:*
               2:   /SIDS8/gcgdata/geneseq/geneseqp/AA[illegible]:*
               3:   /SIDS8/gcgdata/geneseq/geneseqp/AA[illegible]:*
               4:   /SIDS8/gcgdata/geneseq/geneseqp/AA[illegible]:*
               5:   /SIDS8/gcgdata/geneseq/geneseqp/AA[illegible].DAT:*
               6:   /SIDS8/gcgdata/geneseq/geneseqp/AA[illegible]:*
               7:   /SIDS8/gcgdata/geneseq/geneseqp/AA[illegible].DAT:*
               8:   /SIDS8/gcgdata/geneseq/geneseqp/AA[illegible]:*
               9:   /SIDS8/gcgdata/geneseq/geneseqp/AA[illegible]:*
               10:   /SIDS8/gcgdata/geneseq/geneseqp/AA[illegible]:*
               11:   /SIDS8/gcgdata/geneseq/geneseqp/AA[illegible]:*
               12:   /SIDS8/gcgdata/geneseq/geneseqp/AA[illegible]:*
               13:   /SIDS8/gcgdata/geneseq/geneseqp/AA[illegible]:*
               14:   /SIDS8/gcgdata/geneseq/geneseqp/AA[illegible]:*
               15:   /SIDS8/gcgdata/geneseq/geneseqp/AA[illegible]:*
               16:   /SIDS8/gcgdata/geneseq/geneseqp/AA[illegible]:*
               17:   /SIDS8/gcgdata/geneseq/geneseqp/AA[illegible]:*
               18:   /SIDS8/gcgdata/geneseq/geneseqp/AA[illegible]:*
               19:   /SIDS8/gcgdata/geneseq/geneseqp/AA[illegible]:*
               20:   /SIDS8/gcgdata/geneseq/geneseqp/AA[illegible]:*
               21:   /SIDS8/gcgdata/geneseq/geneseqp/AA[illegible]:*
               22:   /SIDS8/gcgdata/geneseq/geneseqp/AA[illegible]:*

  Pred. No. is the number of results predicted by chance to have a
  score greater than or equal to the score of the result being printed,
  and is derived by analysis of the total score distribution.

                              SUMMARIES
                   %
Result             Query
   No.    Score    Match Length DB  ID                        Description
-------------------------------------------------------------------------------
     1       66    100.0     12  20  AAW[illegible]          Mouse steroidogeni
     2       66    100.0     12  22  AAY[illegible]          Mouse [illegible]
     3       66    100.0    284  20  AAW[illegible]          Mouse steroidogeni
     4       66    100.0    284  22  AAY[illegible]          Mouse steroidogeni
     5       61     92.4    284  17  AAW[illegible]          Human steroidogen
     6       61     92.4    285  17  AAW[illegible]          Human steroidogene
     7       61     92.4    285  17  AAW[illegible]          Human steroidogene
     8       61     92.4    285  17  AAW00814                Human steroidogene
     9       61     92.4    285  17  AAW[illegible]          Human steroidogene
    10       61     92.4    285  17  AAW[illegible]          Human steroidogene
    11       61     92.4    285  17  AAW[illegible]          Human steroidogene
    12       61     92.4    285  22  AAY72838                Human steroidogeni
    13       52     78.8     11  22  AAY72[illegible]        Mouse steroidogeni
    14       48     72.7    824  21  AAR[illegible]          Bacillus sp. KSM S
    15       43     65.2    800   8  AAP7042[illegible]      Sequence encoded b
    16       43     65.2    822  18  AAR[illegible]          Alkaline cellulase
    17       39     59.1    346  21  AAY84[illegible]        Amino acid sequenc
    18       38     57.6      6  22  AAY[illegible]          Steroidogenic acut
    19       38     57.6    102  20  AAY[illegible]          pk855 1 secreted p
    20       38     57.6    219  20  AAW[illegible]          expressed antigen
    21       38     57.6    [illegible]30  21  AAB[illegible]  Marine sponge sili
    22       38     57.6   1974  19  AAW[illegible]          H. pylori GHP0 57
    23       37     56.1     73  21  AAG03[illegible]        Human secreted pro
    24       37     56.1     81  21  AAG2[illegible]         Zea mays protein t
    25       37     56.1    113  21  AAG44[illegible]        Zea mays protein t
    26       37     56.1    120  21  AAG26998                Zea mays protein t
    27       37     56.1    128  21  AAG[illegible]          Zea mays protein t
    28       37     56.1    145  21  AAG27[illegible]        Zea mays protein t
    29       37     56.1    159  21  AAG44[illegible]        Zea mays protein t
    30       37     56.1    162  21  AAG54[illegible]        Zea mays protein t
    31       37     56.1    165  21  AAG27[illegible]        Zea mays protein t
    32       37     56.1    288  19  AAW60[illegible]        Streptococcus pneu
    33       37     56.1    300  22  AAB[illegible]          Soybean dESI homol
    34       37     56.1    315  20  AAW80[illegible]        Human RIP interact
    35       37     56.1    362  21  AAB1227[illegible]      Human [illegible]EX [illegible] 2042
    36       37     56.1    393  19  AAW37[illegible]        Murine Ena Vasp li
    37       37     56.1    395  21  AAY91[illegible]        Human cytoskeleton
    38       37     56.1    418  20  AAY[illegible]          Human ena/VASP lik
    39       37     56.1    418  21  AAY[illegible]          Human ena/VASP lik
    40       37     56.1    950  21  AA[illegible]           Human [illegible]EX [illegible]
    41       37     56.1   1511  21  AAB[illegible]18[illegible]  Yeast transporter
    42       36     54.5     77  22  AAY[illegible]          Protein encoded by
    43       36     54.5     78  21  AA[illegible]           Zea mays protein t
    44       36     54.5    144  18  AAW[illegible]          Imm[illegible] type F
    45       36     54.5    162  22  AAY14[illegible]        Amino acid sequenc

                              ALIGNMENTS

RESULT 1
AAW74155
ID   AAW74[illegible]
XX
AC   AAW74155;
XX
DT   05-MAY-1999  (first entry)
XX
DE   Mouse steroid[illegible] peptide.
XX
KW   StAR; ster[illegible] syndrome;
KW   steroid [illegible] dependent disorder; [illegible];
KW   lipoid congenital adrenal hyperplasia; infertility; sexual maturation;
KW   androgen-responsive tumours; hypo[illegible]ism; therapy; progesterone level.
XX
OS   Mus sp.
XX
PN   US5872230-A.
XX
PD   16-FEB-1999.
XX
PF   04-NOV-1995;   95US-0538960.
XX
PR   04-NOV-1995;   95US-0538960.
XX
PA   (UYTE-) UNIV TEXAS TECH HEALTH SCI CENT.
XX
PI   Clark BJ, Stocco DM;
XX
DR   WPI; 1999-1[illegible]/14.
XX
PT   New nucleic acid encoding ster[illegible] acute re[illegible]atory protein -
PT   useful for treatment of stero[illegible] dependent disorders such as
PT   lipoid congenital adrenal hyperplasia, infertility and
```

```
[illegible]  androgen resp[illegible] [illegible]
XX
[illegible]  Example 2[illegible] [illegible]
XX
[illegible]  This sequence [illegible] tryptic peptide fragment [illegible]
[illegible]  regulatory protein ([illegible]AR) of the [illegible]
[illegible]  its products or derivatives may be used [illegible]
[illegible]  hybridisation [illegible] and in the treatment [illegible]
[illegible]  disorders such as lipoid congenital adrenal hyperplasia [illegible]
[illegible]  infertility, sexual [illegible]
[illegible]  puberty, McCune-Albright syndrome, adrenal hyp[illegible]
[illegible]  hypogonadism [illegible] StAR [illegible]
[illegible]  progesterone [illegible]
XX
SQ  Sequence  [illegible] AA;

[illegible]y Match [illegible]
[illegible]st Local Similarity [illegible] Pred. No. [illegible]
  Matches  [illegible] Conservative [illegible] Mismatches [illegible]

[illegible]  I [illegible]ENNQ[illegible]W[illegible] 12
      [illegible]
[illegible]  I [illegible] 12


RESULT  2
AAY72833
ID  AAY72833 standard; peptide; [illegible] AA.
XX
AC  AAY72833;
XX
DT  31 MAY 2001  (first entry)
XX
DE  Mouse tryptic peptide #25 for analysis of 30kDa StAR pr[illegible]
XX
KW  Mouse; steroidogenic acute regulatory protein; StAR
KW  gene therapy; steroid hormone dependent disorder; StAR
KW  lipoid congenital adrenal hyperplasia; sexual [illegible]
KW  infertility; androgen responsive [illegible]
KW  hypogonadotropic hypogonadism; McCune-Albright [illegible]
KW  adrenal hypoplasia [illegible]
XX
[illegible]  Mus sp.
XX
[illegible]  US6194[illegible] B1
[illegible]
PD  27 FEB 2001.
XX
[illegible]  [illegible] JUN [illegible]
XX
[illegible]  [illegible] NOV [illegible]
XX
PA  [illegible]UY[illegible] UNIV [illegible]
XX
PI  Stocco DM [illegible]
XX
DR  WPI; 20[illegible]-21[illegible]
XX
PT  Novel nucleic acid encoding [illegible]
PT  useful for treating steroid hormone dependent [illegible]
XX
PS  Example 2[illegible] Claim [illegible]
XX
[illegible]  The invention relates [illegible]
[illegible]  (StAR) and nucleic acid [illegible] encoding these [illegible]
[illegible]  are used for regulating steroid [illegible]
[illegible]  proteins are involved [illegible]
[illegible]  the mitochondria [illegible]
[illegible]  invention are useful in gene therapy for the treatment [illegible]
[illegible]  hormone dependent disorders [illegible]
[illegible]  adrenal hyperplasia (LCAH) infertility [illegible]
[illegible]  androgen resp[illegible]


[illegible]  [illegible] adrenal hypoplasia [illegible] hypogonadotropic
[illegible]  hypogonadism [illegible] preparation [illegible]
[illegible]  of the steroidogenic acute regulatory protein
[illegible]  The present [illegible] tryptic peptide #[illegible] peptide
[illegible]  used in [illegible] sequence analysis [illegible] DNA encoding the
[illegible]  [illegible] StAR protein [illegible] peptide [illegible]
[illegible]  [illegible] of the StAR protein sequence [illegible] #[illegible]
[illegible]  [illegible]
[illegible]  antigenic determinants [illegible] StAR protein
XX
SQ  Sequence  [illegible] AA;

[illegible] Match [illegible]
[illegible] Local Similarity [illegible] N[illegible]
Matches [illegible] Mismatches [illegible]

[illegible]  [illegible]NNQ[illegible]WP[illegible]
      [illegible]
[illegible]  [illegible]


RESULT [illegible]
AAW[illegible]
ID  AAW[illegible] standard; Protein; [illegible] AA.
XX
AC  AAW[illegible]
XX
DT  [illegible] MAY [illegible] (first entry)
XX
DE  M[illegible] steroidogenic acute regulatory protein
XX
KW  StAR; steroidogenic acute regulatory protein; McCune Albright syndrome;
KW  steroid hormone dependent disorder; adrenal hypoplasia congenita; LCAH;
KW  lipoid congenital adrenal hyperplasia; infertility; sexual maturation;
KW  androgen responsive tumours; hypogonadism; therapy; progesterone level.
XX
[illegible]  Mus sp.
XX
PN  [illegible]
XX
[illegible]  [illegible]
XX
[illegible]  [illegible] NOV [illegible]
XX
[illegible]  [illegible] NOV [illegible]
XX
PA  [illegible] UNIV [illegible] TEXAS [illegible]
XX
PI  Stocco [illegible]
XX
DR  WPI; [illegible]
DR  N-PSDB; AAT[illegible]
XX
PT  Novel [illegible] steroidogenic acute regulatory protein [illegible]
PT  useful for treatment of steroid hormone dependent disorders such as
PT  [illegible]
PT  androgen responsive tumours [illegible]
XX
PS  Claim [illegible]
XX
[illegible]  This sequence [illegible] steroidogenic acute regulatory protein (StAR)
[illegible]  [illegible] StAR [illegible]
[illegible]  derivatives may be used in probes and primers for PCR hybridisation
[illegible]  studies and in the treatment of steroid hormone dependent disorders such
[illegible]  as lipoid congenital adrenal hyperplasia (LCAH), infertility, sexual
[illegible]  maturation, androgen responsive tumours [illegible] puberty,
[illegible]  McCune-Albright syndrome, adrenal hypoplasia congenita and hypogonadism
[illegible]  [illegible] StAR [illegible]
[illegible]  and [illegible]
XX
SQ  Sequence  [illegible] AA;
```

```
  Query Match             100.0%;  Score 66;  DB 20;  Length 284;
  Best Local Similarity   100.0%;  Pred. No. 0.00095;
  Matches   12;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ALGILNNQEGWK 12
              ||||||||||||
Db         85 alqilnnqeqwk 96


RESULT    4
AAY72831
ID   AAY72831 standard; Protein; 284 AA.
XX
AC   AAY72831;
XX
DT   31-MAY-2001  (first entry)
XX
DE   Mouse steroidogenic acute regulatory (StAR) protein.
XX
KW   Mouse; steroidogenic acute regulatory protein; StAR;
KW   gene therapy; steroid hormone-dependent disorder; LCAH;
[illegible]   lipoid congenital adrenal hyperplasia; sexual maturation;
[illegible]   infertility; androgen-responsive tumour; precocious puberty;
[illegible]   hypogonadotropic hypogonadism; McCune Albright syndrome;
KW   adrenal-hypoplasia congenita; cytostatic.
XX
OS   Mus sp.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..47
FT                   /label= Signal_peptide
FT   Cleavage-site   39..40
FT                   /note= "Site for mitochondrial proteases,
FT                   matrix-processing protease and mitochondrial
FT                   intermediate processing peptide"
FT   Protein         48..284
FT                   /label= Mature_steroidogenic_acute_regulatory_protein
FT   Cleavage-site   49..50
FT                   /note= "Site for mitochondrial proteases,
FT                   matrix processing protease and mitochondrial
FT                   intermediate processing peptide"
FT   Modified-site   59..60
FT                   /note= "PKA phosphorylation site"
FT   Modified-site   68..69
FT                   /note= "CK phosphorylation site"
FT   Region          85..96
FT                   /label= Tryptic_peptide_#25
FT   Region          88..98
FT                   /note= "This sequence is used to raise antibodies
[illegible]                   against StAR protein"
[illegible]   Region          193..211
[illegible]                   /label= Tryptic_peptide_#45
FT   Modified-site   195
FT                   /note= "PKA phosphorylation site"
FT   Region          217..230
FT                   /label= Tryptic_peptide_#23
FT   Modified-site   233..234
FT                   /note= "P34CDC2 kinase (Cd) phosphorylation site"
FT   Modified-site   238
FT                   /note= "PKC phosphorylation site"
XX
PN   US6194555 B1.
XX
PD   27-FEB-2001.
XX
PF   07-JUN-1996;   96US-0659254.
XX
PR   04-NOV-1995;   95US-0538960.
XX
PA   (UYTE-) UNIV TEXAS TECH HEALTH SCI CENT.
XX
PI   Stocco DM,  Clark BJ;
XX
DR   WPI; 2001-217936/22.
DR   N-PSDB; AAB02901, AAB02902.
XX
PT   Novel nucleic acid encoding steroidogenic acute regulatory protein,
PT   useful for treating steroid hormone-dependent disorders
XX
PS   Example 2; Fig 2; 41pp; English.
XX
CC   The invention relates to steroidogenic acute regulatory protein
CC   (StAR) and nucleic acid molecules encoding them. StAR proteins
CC   are used for regulating steroidogenesis. More particularly these
CC   proteins are involved in regulating cholesterol transport into
CC   the mitochondria of a cell. The nucleic acid molecules of the
CC   invention are useful in gene therapy for the treatment of steroid
CC   hormone-dependent disorders, particularly, lipoid congenital
CC   adrenal hyperplasia (LCAH), infertility, sexual maturation,
CC   androgen-responsive tumours, precocious puberty, McCune Albright
CC   syndrome, adrenal hypoplasia congenita or hypogonadotropic
CC   hypogonadism. They are also useful in recombinant preparation
CC   of the steroidogenic acute regulatory protein.
CC   The present sequence is mouse steroidogenic acute regulatory
CC   protein (StAR).
XX
SQ   Sequence   284 AA;


  Query Match             100.0%;  Score 66;  DB 22;  Length 284;
  Best Local Similarity   100.0%;  Pred. No. 0.00095;
  Matches   12;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ALGILNNQEGWK 12
              ||||||||||||
Db         85 alqilnnqeqwk 96


RESULT    5
AAW13238
ID   AAW13238 standard; Protein; 284 AA.
XX
AC   AAW13238;
XX
DT   21-MAY-1997  (first entry)
XX
DE   Human steroidogenesis acute regulatory protein mutant delta Arg272.
XX
KW   Human; steroidogenesis; acute regulatory protein; hStAR; analysis;
KW   mutation; detection; prenatal; genetic defect; congenital; protein;
KW   lipoid adrenal hyperplasia; treatment; prevention; gene;
KW   replacement therapy; hypercholesterolaemia; mutant.
XX
OS   Homo sapiens.
XX
PN   WO9629338-A1.
XX
PD   26-SEP-1996.
XX
PF   22-MAR-1996;   96WO-US03896.
XX
PR   23-MAR-1995;   95US-0410540.
XX
PA   (REGC ) UNIV CALIFORNIA.
PA   (UYPE ) UNIV PENNSYLVANIA.
XX
PI   Lin D,  Miller WL,  Strauss JF;
XX
DR   WPI; 1996-443139/44.
XX
PT   Isolated human steroidogenesis acute regulatory protein gene - used
PT   for detection of mutation(s) of this gene that cause congenital
PT   lipoid adrenal hyperplasia
XX
PS   Claim 16; Page ; 89pp; English.
```

```
XX
CC   The present sequence is the human steroidogenic [illegible]
CC   protein (hSTAR) [illegible] lipoid adrenal [illegible]
CC   associated mutant [illegible] Arg[illegible] ([illegible] wild type [illegible] Arg [illegible]
CC   been deleted). Analysis of the hSTAR gene [illegible]
CC   useful in the detection (e.g. prenatally) of [illegible]
CC   genetic detection or its transmission to children. [illegible]AH [illegible]
CC   treated by protein or gene replacement therapy [illegible]
CC   used to prevent or treat hypercholesterolaemia.
XX
SQ   Sequence [illegible] AA;

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches [illegible]; Conservative [illegible]; Mismatches [illegible]

Qy        1 A[illegible]LNN[illegible]EWR[illegible]
              [illegible]
Db       86 [illegible]

RESULT 6
AAW13286
ID   AAW13286 standard; Protein; [illegible] AA.
XX
AC   AAW13286;
XX
DT   21 MAY 1997 (first entry)
XX
DE   Human steroidogenesis acute regulatory protein [illegible]
XX
KW   Human; steroidogenesis acute regulatory protein; [illegible]
KW   mutation; detection; prenatal; genetic defect; [illegible]
KW   lipoid adrenal hyperplasia; treatment; prevention; [illegible]
KW   replacement therapy; hypercholesterolaemia; [illegible]
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc_difference 238
FT                   /note= "wild type A[illegible]"
XX
PN   W0962[illegible]-A1.
XX
PD   26 SEP 1996.
XX
PF   22 MAR 1996;   [illegible]
XX
PR   23 MAR 1995;   [illegible]
XX
PA   (REGC ) UNIV CALIFORNIA.
PA   (UYPE ) UNIV PENNSYLVANIA.
XX
PI   Lin D, Miller WL, Strauss JF;
XX
DR   WPI; 1996-44[illegible].
DR   N-PSDB; AAT[illegible].
XX
PT   Isolated human steroidogenesis acute regulatory protein [illegible]
PT   for detection [illegible]
PT   lipoid adrenal hyperplasia
XX
PS   Claim [illegible]; Page [illegible]
XX
CC   The present sequence [illegible]
CC   protein (hSTAR), [illegible] lipoid adrenal hyperplasia [illegible]
CC   associated mutant A[illegible]V[illegible]. Analysis [illegible]
CC   mutation is useful in the detection [illegible]
CC   associated genetic detection [illegible]
CC   can be treated by protein [illegible]
CC   also be used to prevent or treat hypercholesterolaemia.
XX
```

```
SQ   Sequence [illegible] AA;

  Query Match [illegible]; Score [illegible]; DB [illegible]; Length [illegible];
  Best Local Similarity [illegible]; Pred. No. [illegible];
  Matches [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible];

Qy        [illegible] A[illegible]NN[illegible]EWR[illegible]
              [illegible]
Db        [illegible]

RESULT [illegible]
AAW[illegible]
ID   AAW[illegible] standard; Protein; [illegible] AA.
XX
AC   AAW[illegible];
XX
DT   [illegible] MAY [illegible] (first entry)
XX
DE   Human steroidogenesis acute regulatory protein mutant [illegible]
XX
KW   Human; steroidogenesis acute regulatory protein; hSTAR; [illegible]
KW   mutation; detection; prenatal; genetic defect; congenital; [illegible]
KW   lipoid adrenal hyperplasia; treatment; prevention; [illegible]
KW   replacement therapy; hypercholesterolaemia; [illegible]
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc_difference [illegible]
FT                   /note= "wild type [illegible] mutated with Pro"
XX
PN   W[illegible]-A1.
XX
PD   [illegible] SEP [illegible]
XX
PF   [illegible] MAR [illegible]
XX
PR   [illegible] MAR [illegible]
XX
PA   ([illegible]) UNIV CALIFORNIA.
PA   ([illegible]) UNIV PENNSYLVANIA.
XX
PI   [illegible]
XX
DR   WPI; [illegible]
DR   N-PSDB; AA[illegible]
XX
PT   Isolated human [illegible]
PT   [illegible]
PT   lipoid adrenal hyperplasia
XX
PS   [illegible]
XX
CC   The present sequence [illegible]
CC   protein (hSTAR) [illegible] lipoid adrenal hyperplasia (LAH)
CC   [illegible]. Analysis of the hSTAR gene [illegible]
CC   [illegible] the detection (e.g. prenatally) of LAH
CC   [illegible] transmission [illegible] LAH
CC   can be treated by [illegible] gene replacement therapy, which can
CC   also be used to prevent or treat hypercholesterolaemia.
XX
SQ   Sequence [illegible] AA;

  Query Match [illegible]; Length [illegible];
  Best Local Similarity [illegible]; Pred. No. [illegible];
  Matches [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible];

Qy        [illegible]NN[illegible]
              [illegible]
Db        [illegible]
```

```
RESULT   8
AAW00814
ID   AAW00814 standard; Protein; 285 AA
XX
AC   AAW00814;
XX
DT   21-MAY-1997  (first entry)
XX
DE   Human steroidogenesis acute regulator, protein.
XX
KW   Human; steroidogenesis; acute regulatory protein; hStAR; analysis;
KW   mutation; detection; prenatal; genetic defect; congenital; protein;
KW   lipoid adrenal hyperplasia; treatment; prevention; gene;
KW   replacement therapy; hypercholesterolaemia.
XX
OS   Homo sapiens
XX
FH   Key                 Location/Qualifiers
FT   Region              5
FT                       /note= "potential site for protein kinase C
FT                       mediated phosphorylation"
     Region              13
                         /note= "potential site for protein kinase C
FT                       mediated phosphorylation"
FT   Region              56..57
FT                       /note= "potential site for protein kinase A
FT                       mediated phosphorylation"
FT   Region              100
FT                       /note= "potential site for protein kinase A
FT                       mediated phosphorylation"
FT   Region              110
FT                       /note= "potential site for protein kinase A
FT                       mediated phosphorylation"
FT   Region              186
FT                       /note= "potential site for protein kinase C
FT                       mediated phosphorylation"
FT   Region              195..196
FT                       /note= "potential site for protein kinase A
FT                       mediated phosphorylation"
FT   Region              277
FT                       /note= "potential site for protein kinase A
FT                       mediated phosphorylation"
XX
PN   WO9629338-A1.
XX
PD   26-SEP-1996.
XX
PF   22-MAR-1996;   96WO-US03896.
XX
     23-MAR-1995;   95US-0410540.
XX
PA   (REGC ) UNIV CALIFORNIA.
PA   (UYPE-) UNIV PENNSYLVANIA.
XX
PI   Lin D,  Miller WL,  Strauss JF;
XX
DR   WPI; 1996-443130/44.
DR   N-PSDB; AAT33650.
XX
PT   Isolated human steroidogenesis acute regulatory protein gene - used
PT   for detection of mutation(s) of this gene that cause congenital
PT   lipoid adrenal hyperplasia
XX
PS   Claim 1; Pages 62-63; 89pp; English.
XX
CC   The present sequence is the human steroidogenesis acute
CC   regulatory protein (hStAR). The hStAR gene can be analysed for
CC   mutations to detect (e.g. prenatally) genetic defects associated
CC   with congenital lipoid adrenal hyperplasia (CAH), or its
CC   transmission to children. CAH can be treated by protein or gene
CC   replacement therapy, which can also be used to prevent or treat
CC   hypercholesterolaemia.
CC   A human adrenal cortex cDNA library was screened with a mouse StAR
CC   probe to isolate a 1.6 kb insert, encoding an ORF for a 285
CC   residue protein. When it was cloned into pSPKF1 and expressed in
CC   COS-1 cells cotransfected with pP450scc and pADX, it increased the
CC   level of pregnenolone synthesis from cholesterol
CC   20-alpha-hydroxycholesterol.
XX
SQ   Sequence   285 AA;


  Query Match             92.4%;  Score 61;  DB 17;  Length 285;
  Best Local Similarity   91.7%;  Pred. No. [illegible];
  Matches   11;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 ALGILNNQEGWK 12
            |||||:||| ||
Db       86 aiqilsnqeqwk 97


RESULT   9
AAW13233
ID   AAW13233 standard; Protein; 285 AA.
XX
AC   AAW13233;
XX
DT   21 MAY 1997  (first entry)
XX
DE   Human steroidogenesis acute regulatory protein mutant Glu169Gly.
XX
KW   Human; steroidogenesis; acute regulatory protein, hStAR; analysis;
KW   mutation; detection; prenatal; genetic defect; congenital; protein;
KW   lipoid adrenal hyperplasia; treatment; prevention; gene;
KW   replacement therapy; hypercholesterolaemia; mutant
XX
OS   Homo sapiens.
XX
FH   Key                 Location/Qualifiers
FT   Misc-difference 169
FT                       /note "wild type Glu substituted with Gly"
XX
PN   WO9629338-A1.
XX
PD   26-SEP-1996.
XX
PF   22-MAR-1996;   96WO-US03896.
XX
PR   23 MAR 1995;   95US 0410540.
XX
PA   (REGC ) UNIV CALIFORNIA.
PA   (UYPE-) UNIV PENNSYLVANIA.
XX
PI   Lin D,  Miller WL,  Strauss JF;
XX
DR   WPI; 1996-443130/44.
DR   N-PSDB; AAT62013.
XX
PT   Isolated human steroidogenesis acute regulatory protein gene - used
PT   for detection of mutation(s) of this gene that cause congenital
PT   lipoid adrenal hyperplasia
XX
PS   Claim 13; Page  ; 89pp; English.
XX
CC   The present sequence is the human steroidogenesis acute regulatory
CC   protein (hStAR), congenital lipoid adrenal hyperplasia (CAH)
CC   associated mutant, Glu169Gly. Analysing the hStAR gene for such a
CC   mutation is useful in the detection (e.g. prenatally) of CAH
CC   associated genetic defects, or its transmission to children. CAH
CC   can be treated by protein or gene replacement therapy, which can
CC   also be used to prevent or treat hypercholesterolaemia.
XX
SQ   Sequence   285 AA;
```

```
  Query Match             [illegible]
  Best Local Similarity   [illegible]
  Matches   11;  [illegible]  Mismatches  [illegible]

Qy          1 ALGILNNLEGWK 1[illegible]
              [illegible]
Db         86 alqilsnqeswk [illegible]


RESULT 10
AAW13234
ID   AAW13234 standard; protein; 28[illegible] AA.
XX
AC   AAW13234;
XX
DT   21 MAY 1997 [illegible]
XX
DE   Human steroidogenesis acute regulatory protein [illegible]
XX
KW   Human; steroidogenesis acute regulatory protein; [illegible]
KW   mutation; detection; prenatal; genetic defect; [illegible]
KW   lipoid adrenal hyperplasia; treatment; prevent[illegible]
KW   replacement therapy; hypercholesterolaemia[illegible]
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 192
FT                   /note= "wild type Arg [illegible]"
XX
PN   WO962[illegible] A1.
XX
PD   26-SEP-1996.
XX
PF   22-MAR-1996;   96WO-[illegible]
XX
PR   23-MAR-1995;   95US-[illegible]540.
XX
PA   (REGC ) UNIV CALIFORNIA.
PA   (UYPE ) UNIV PENNSYLVANIA.
XX
PI   Lin D,  Miller W[illegible]  Strauss JF;
XX
DR   WPI; 1996-441[illegible]/44.
DR   N-PSDB; AAT62[illegible]
XX
PT   Isolated human steroidogenesis acute regulatory protein [illegible]
PT   for detection of mutation(s) of this gene that [illegible]
PT   lipoid adrenal hyperplasia.
XX
PS   Claim 1; Page [illegible]; 99pp; English.
XX
CC   The present sequence is the human steroidogenesis acute regulatory
CC   protein (hSTAR). Congenital lipoid adrenal hyperplasia (CAH)
CC   associated mutant [illegible]. Analysing the hSTAR gene for such a
CC   mutation is useful in the detection (e.g. prenatally) of CAH
CC   associated genetic defects, or its transmission to children. CAH
CC   can be treated by protein or gene replacement therapy, which can
CC   also be used to prevent or treat hypercholesterolaemia.
XX
SQ   Sequence   285 AA;

  Query Match             [illegible]
  Best Local Similarity   [illegible]
  Matches   11;  Conservative  [illegible]  Mismatches  [illegible]  Indels  [illegible]  Gaps  [illegible]

Qy          1 ALGILNNLEGWK 12
              [illegible]
Db         86 alqilsnqeswk 97
```

```
RESULT 11
AAW[illegible]
ID   AAW[illegible] standard; protein; 285 AA.
XX
AC   AAW[illegible]
XX
DT   [illegible] MAY [illegible]
XX
DE   [illegible] steroidogenesis acute regulatory protein [illegible] mutant [illegible]
XX
KW   [illegible] steroidogenesis acute regulatory protein; hSTAR; [illegible]
KW   mutation; detection; prenatal; genetic defect; congenital; [illegible]
KW   lipoid adrenal hyperplasia; treatment; [illegible]
KW   replacement therapy; hypercholesterolaemia; [illegible]
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference [illegible]
FT                   /note= "wild type [illegible] with [illegible]"
XX
PN   W[illegible] A[illegible]
XX
PD   [illegible]
XX
PF   [illegible] MAR [illegible]
XX
PR   [illegible] MAR [illegible] 4[illegible]
XX
PA   (REGC ) UNIV CALIFORNIA
PA   (UYPE ) UNIV PENNSYLVANIA
XX
PI   [illegible] W[illegible]
XX
DR   W[illegible] 44[illegible]/44.
DR   N-PSDB; AA[illegible]
XX
PT   Isolated human steroidogenesis acute regulatory protein [illegible]
PT   for detection of mutation(s) of this gene that cause congenital
PT   lipoid adrenal hyperplasia.
XX
PS   Claim [illegible]; Page [illegible]; 99pp; English.
XX
CC   The present sequence is the human steroidogenesis acute regulatory
CC   protein (hSTAR). Congenital lipoid adrenal hyperplasia (CAH)
CC   associated mutant [illegible]. Analysing the hSTAR gene for such a
CC   mutation is useful in the detection (e.g. prenatally) of CAH
CC   associated genetic defects, or its transmission to children. CAH
CC   can be treated by protein or gene replacement therapy, which can
CC   also be used to prevent or treat hypercholesterolaemia.
XX
SQ   Sequence   285 AA;

  Query Match             [illegible]
  Best Local Similarity   [illegible]  Pred. No.  [illegible]
  Matches   [illegible]  Conservative  [illegible]  Mismatches  [illegible]  Indels  [illegible]  Gaps  [illegible]

Qy          1 A[illegible]NNLEGWK [illegible]

Db         [illegible] qeswk [illegible]


RESULT 12
AAY2[illegible]
ID   AAY2[illegible] standard; protein; 2[illegible] AA.
XX
AC   AAY2[illegible]
XX
DT   [illegible] MAY [illegible]
XX
DE   Human steroidogenic acute regulatory (STAR) protein [illegible]
XX
```

```
KW   Human; steroidogenic acute regulatory protein; StAR;
KW   gene therapy; steroid hormone-dependent disorder; LCAH;
KW   lipoid congenital adrenal hyperplasia; sexual maturation;
KW   infertility; androgen-responsive tumour; precocious puberty;
KW   hypogonadotropic hypogonadism; McCune Albright syndrome;
KW   adrenal-hypoplasia congenita; cytostatic
XX
OS   Homo sapiens.
XX
FH   Key                 Location/Qualifiers
FT   Misc-difference     156
FT                       /note  "Encoded by GAT; this feature corresponds
FT                       to the translation exception for the sequence shown
FT                       in (AAD02909)"
FT   Modified-site       59..60
FT                       /note  "PKA phosphorylation site"
FT   Modified-site       68..69
FT                       /note  "CK phosphorylation site"
FT   Modified-site       195
FT                       /note  "PKA phosphorylation site"
FT   Modified-site       233..234
FT                       /note  "P34CDC2 kinase (Pk) phosphorylation site"
     Modified-site       238
                         /note  "PKC phosphorylation site"
F    Modified-site       277
FT                       /note  "PKA phosphorylation site"
XX
PN   US6194555-B1.
XX
PD   27-FEB-2001.
XX
PF   07-JUN-1996;   96US-0659254.
XX
PR   04-NOV-1995;   95US-0538960.
XX
PA   (UYTE-) UNIV TEXAS TECH HEALTH SC CENT.
XX
PI   Stocco DM,  Clark BJ;
XX
DR   WPI; 2001 217936/22.
DR   N-PSDB; AAD02909, AAD02912.
XX
PT   Novel nucleic acid encoding steroidogenic acute regulatory protein,
PT   useful for treating steroid hormone-dependent disorders -
XX
PS   Example 14; Fig 5; 41pp; English.
XX
CC   The invention relates to steroidogenic acute regulatory protein
CC   (StAR) and nucleic acid molecules encoding them. StAR proteins
CC   are used for regulating steroidogenesis. More particularly these
CC   proteins are involved in regulating cholesterol transport into
     the mitochondria of a cell. The nucleic acid molecules of the
     invention are useful in gene therapy for the treatment of steroid
CC   hormone-dependent disorders, particularly, lipoid congenital
CC   adrenal hyperplasia (LCAH), infertility, sexual maturation,
CC   androgen-responsive tumours, precocious puberty, McCune-Albright
CC   syndrome, adrenal-hypoplasia congenita or hypogonadotropic
CC   hypogonadism. They are also useful for recombinant preparation
CC   of the steroidogenic acute regulatory protein.
CC   The present sequence is human steroidogenic acute regulatory
CC   protein (StAR).
XX
SQ   Sequence   285 AA;

  Query Match             92.4%;  Score [illegible];  DB 22;  Length 285;
  Best Local Similarity   91.7%;  Pred. No. [illegible];
  Matches   11;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ALGILNNQEGWK 12
                ||:||:||||||
Db         86 alqilsnqeqwk 97
```

```
RESULT 13
AAY72837
ID   AAY72837 standard; peptide; 11 AA.
XX
AC   AAY72837;
XX
DT   31-MAY-2001  (first entry)
XX
DE   Mouse steroidogenic acute regulatory (StAR) peptide #4.
XX
KW   Mouse; steroidogenic acute regulatory protein; StAR;
KW   gene therapy; steroid hormone-dependent disorder; LCAH;
KW   lipoid congenital adrenal hyperplasia; sexual maturation;
KW   infertility; androgen responsive tumour; precocious puberty;
KW   hypogonadotropic hypogonadism; McCune Albright syndrome;
KW   adrenal-hypoplasia congenita; cytostatic.
XX
OS   Mus sp.
XX
PN   US6194555-B1.
XX
PD   27-FEB-2001.
XX
PF   07-JUN-1996;   96US-0659254.
XX
PR   04-NOV-1995;   95US-0538960.
XX
PA   (UYTE-) UNIV TEXAS TECH HEALTH SC CENT.
XX
PI   Stocco DM,  Clark BJ;
XX
DR   WPI; 2001 217936/22.
XX
PT   Novel nucleic acid encoding steroidogenic acute regulatory protein,
PT   useful for treating steroid hormone dependent disorders -
XX
PS   Disclosure; Column 51; 41pp; English.
XX
CC   The invention relates to steroidogenic acute regulatory protein
CC   (StAR) and nucleic acid molecules encoding them. StAR proteins
CC   are used for regulating steroidogenesis. More particularly these
CC   proteins are involved in regulating cholesterol transport into
CC   the mitochondria of a cell. The nucleic acid molecules of the
CC   invention are useful in gene therapy for the treatment of steroid
CC   hormone-dependent disorders, particularly, lipoid congenital
CC   adrenal hyperplasia (LCAH), infertility, sexual maturation,
CC   androgen responsive tumours, precocious puberty, McCune Albright
CC   syndrome, adrenal-hypoplasia congenita or hypogonadotropic
CC   hypogonadism. They are also useful for recombinant preparation
CC   of the steroidogenic acute regulatory protein.
CC   The present sequence is steroidogenic acute regulatory (StAR)
CC   peptide. This peptide corresponds to amino acids 88-98 of mouse
CC   StAR protein. This peptide is used for design of oligonucleotides
CC   for screening and identifying antigenic determinants of StAR protein.
XX
SQ   Sequence   11 AA;

  Query Match             78.8%;  Score 52;  DB 22;  Length 11;
  Best Local Similarity   100.0%;  Pred. No. [illegible];
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 ILNNQEGWK 12
                |||||| |
Db          1 ilnnqeqwk 9


RESULT 14
AAB23180
ID   AAB23180 standard; Protein; 824 AA
XX
AC   AAB23180;
```

```
XX
DT   29-JAN-2001  (first entry)
XX
DE   Bacillus sp. KSM-S2[illegible] heat re[illegible]
XX
KW   Heat resistant alkaline cellulase; detergent [illegible]
XX
OS   Bacillus sp. KSM-S2[illegible]
XX
PN   JP2000210081-A.
XX
PD   02-AUG-2000.
XX
PF   21-JAN-1999;   99JP-[illegible]
XX
PR   21-JAN-1999;   99JP-[illegible]
XX
PA   (KAOS ) KAO [illegible]
XX
DR   WPI; 2000-[illegible]
DR   N-PSDB; AAA[illegible]
XX
PT   A heat resistant alkaline cellulase gene.
XX
PS   Claim 1; Page [illegible]; Japanese.
XX
CC   The invention relates to [illegible] heat [illegible]
CC   (AAB23186) from Bacillus sp. KSM-S23[illegible]
CC   (AAA97403). The invention also [illegible]
CC   resistant alkaline cellulase in which at least [illegible]
CC   deleted, replaced or inserted. The Bacillus sp. KSM-[illegible]
CC   alkaline cellulase gene was isolated [illegible]
CC   alkaline cellulase gene derived PCR primers AAA[illegible]
CC   sp. KSM-S237 heat resistant alkaline [illegible]
CC   recombinant production of heat resistant alkaline [illegible]
CC   formulated in detergent [illegible]
CC   represents the Bacillus sp. KSM-[illegible]
XX
SQ   Sequence [illegible] AA;

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches [illegible] Conservative [illegible] Mismatches [illegible]

Qy [illegible] NNQF[illegible]WK
              [illegible]
Db       209 [illegible]pnneewk [illegible]


RESULT 15
AAP70420
ID   AAP70420 standard; protein; 8[illegible] AA.
XX
AC   AAP70420;
XX
DT   20-JAN-1991  (first entry)
XX
DE   Sequence encoded by cellulase gene derived from [illegible]
XX
KW   Enzyme; cellulose; [illegible]
XX
OS   Bacillus sp. N[illegible]
XX
FH   Key                [illegible]
FT   Peptide            [illegible]
FT   Protein            [illegible]
XX
PN   JP62232086-A.
XX
PD   12-OCT-1987.
XX
PF   02-APR-1987;   [illegible]
XX
PR   [illegible] APR [illegible]
XX
PA   [illegible] KA[illegible] [illegible]NKY [illegible]
XX
DR   W[illegible]
DR   N-PSDB; AAN[illegible]
XX
PT   DNA sequence coding for [illegible] gene [illegible] derived from Bacillus sp. [illegible]
PT   [illegible]
XX
PS   [illegible] Japanese.
XX
CC   [illegible]
CC   [illegible]
CC   [illegible]
XX
SQ   Sequence [illegible] AA;

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches [illegible] Conservative [illegible] Mismatches [illegible] Indels [illegible] Gaps [illegible]

Qy [illegible] NNQ[illegible]W[illegible]
              [illegible]
Db [illegible]


Search completed: August [illegible]
[illegible]
```

```
                        GenCore version 1.5
              Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on:          August 22, 2001, 14:30:10 ; Search time 55.29 Seconds
                                            (without alignments)
                                            4.469 Million cell updates/sec

Title:           US-08-659-254-4
Perfect score:   66
Sequence:        1 ALGILNNQEGWK 12

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        197339 seqs, 20590346 residues

Total number of hits satisfying chosen parameters:      197339

[illegible]mum DB seq length: 0
[illegible]mum DB seq length: 2000000000

Post processing  Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :        Issued_Patents_AA:*
               1:  /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
               2:  /cgn2_6/ptodata/2/iaa/[illegible]_COMB.pep:*
               3:  /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
               4:  /cgn2_6/ptodata/2/iaa/[illegible]_COMB.pep:*
               5:  /cgn2_6/ptodata/2/iaa/[illegible]_COMB.pep:*
               6:  /cgn2_6/ptodata/2/iaa/[illegible].pep:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed
     and is derived by analysis of the total score distribution.

                              SUMMARIES

                %
Result          Query
  No.   Score   Match Length DB  ID                              Description
-------------------------------------------------------------------------------
    1      66   100.0     12  2  US-08-538-960-4                 Sequence 4, Appl
    2      66   100.0     12  4  US-08-659-254-4                 Sequence 4, Appl
    3      66   100.0    276  4  US-08-659-254-2                 Sequence 2, Appl
    4      66   100.0    284  2  US-08-538-960-2                 Sequence 2, Appl
    5      61    92.4    285  1  US-08-41[illegible]-2            Sequence 2, Appl
    6      61    92.4    285  4  US-08-659-254-18                Sequence 18, Appl
    7      52    78.8     11  2  US-08-538-960-8                 Sequence 8, Appl
    8      52    78.8     11  4  US-08-659-254-8                 Sequence 8, Appl
    9      38    57.6      6  2  US-08-538-960-6                 Sequence 6, Appl
   10      38    57.6      6  4  US-08-659-254-6                 Sequence 6, Appl
   11      37    56.1    393  2  US-09-02[illegible]-3           Sequence 3, Appl
   12      37    56.1    393  2  US-09-227-420-3                 Sequence 3, Appl
   13      37    56.1    418  2  US-09-02[illegible]-87-1        Sequence 1, Appl
   14      37    56.1    418  2  US-09-227-420-1                 Sequence 1, Appl
   15      36    54.5    161  2  US-08-997-[illegible]0-158      Sequence 158, App
   16      36    54.5    161  2  US-08-997-[illegible]2-158      Sequence 158, App
   17      36    54.5    161  4  US-09-09[illegible]5-158        Sequence 158, App
   18      36    54.5    308  3  US-08-50[illegible]-35          Sequence 35, Appl
   19      36    54.5    326  2  US-08-997-[illegible]80-43      Sequence 43, Appl
   20      36    54.5    326  2  US-08-997-[illegible]62-43      Sequence 43, Appl
   21      36    54.5    326  3  US-08-97[illegible]0-43         Sequence 43, Appl
   22      36    54.5    326  4  US-0[illegible]55-43            Sequence 43, Appl
   23      36    54.5    330  2  US-08-[illegible]26-1           Sequence 1, Appl
   24      36    54.5    333  2  US-08-997-[illegible]80-36      Sequence 36, Appl
   25      36    54.5    333  2  US-08-997-[illegible]62-36      Sequence 36, Appl
   26      36    54.5    333  3  US-08-87[illegible]0-36         Sequence 36, Appl
   27      36    54.5    333  4  US-09-0[illegible]55-36         Sequence 36, Appl
   28      36    54.5    334  2  US-08-997-[illegible]80-45      Sequence 45, Appl
   29      36    54.5    334  2  US-08-997-362-45                Sequence 45, Appl
   30      36    54.5    334  3  US-08-8[illegible]-45           Sequence 45, Appl
   31      36    54.5    334  4  US-09-035-85[illegible]-45      Sequence 45, Appl
   32      36    54.5    340  2  US-08-[illegible]               Sequence 3, Appli
   33      36    54.5    340  2  US-08-[illegible]               Sequence 37, Appl
   34      36    54.5    340  2  US-08-997-080-37                Sequence 37, Appl
   35      36    54.5    340  2  US-08-997-362-37                Sequence 37, Appl
   36      36    54.5    340  3  US-08-[illegible]9[illegible]-37 Sequence 37, Appl
   37      36    54.5    340  4  US-09-[illegible]5-85[illegible]-37 Sequence 37, Appl
   38      35    53.0    161  2  US-08-997-080-44                Sequence 44, Appl
   39      35    53.0    161  2  US-08-997-08[illegible]         Sequence 7[illegible], Appl
   40      35    53.0    161  2  US-08-997-36[illegible]-44      Sequence 44, Appl
   41      35    53.0    161  2  US-08-[illegible]36[illegible]  Sequence 7[illegible], Appl
   42      35    53.0    161  3  US-08-87[illegible]-9[illegible]-44 Sequence 44, Appl
   43      35    53.0    161  3  US-08-87[illegible]             Sequence 7[illegible], Appl
   44      35    53.0    161  4  US-09-0[illegible]8[illegible]-44 Sequence 44, Appl
   45      35    53.0    161  4  US-09-0[illegible]8[illegible]  Sequence 7[illegible], Appl
```

```
                              ALIGNMENTS

RESULT 1
US-08-538-960-4
; Sequence 4, Application US/08538960
; Patent No. 5872230
;  GENERAL INFORMATION:
;     APPLICANT:  Stocco, Douglas M.
;     APPLICANT:  Clark, Barbara J.
;     TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR
;     TITLE OF INVENTION:  REGULATION OF STEROIDOGENESIS
;     NUMBER OF SEQUENCES:  14
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Akin, Gump, Strauss, Hauer & Feld, L.L.P.
;       STREET:  [illegible] Park Place, [illegible] Avenue
;       CITY:  Austin
;       STATE:  TX
;       COUNTRY:  U.S.A.
;       ZIP:  78701
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/538,960
;       FILING DATE:  Concurrently Herewith
;       CLASSIFICATION:  530
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Mayfield, Denise L.
;       REGISTRATION NUMBER:  33,732
;       REFERENCE/DOCKET NUMBER:  4[illegible]2/COM
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  512/499-6200
;       TELEFAX:  512/499-6290
;  INFORMATION FOR SEQ ID NO:  4:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  12 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
US-08-538-960-4

  Query Match             100.0%;  Score 66;  DB 2;  Length 12;
  Best Local Similarity   100.0%;  Pred. No. 6.[illegible];
  Matches   12;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ALGILNNQEGWK 12
              ||||||||||||
Db          1 ALGILNNQEGWK 12
```

```
RESULT 2
US-08-659-254-4
; Sequence 1, Application US/08659254
; Patent No. 6194[illegible]
; GENERAL INFORMATION:
;   APPLICANT: [illegible], [illegible] M
;   APPLICANT: Clark, [illegible] Barbara [illegible]
;   TITLE OF INVENTION: COMPOSITIONS AND METHODS [illegible]
;   TITLE OF INVENTION: [illegible] GENE[illegible]
;   NUMBER OF SEQUENCES: [illegible]
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Akin, Gump, Strauss, Hauer & Feld, [illegible]
;     STREET: 1900 Frost Bank Plaza, [illegible] Congress [illegible]
;     CITY: Austin
;     STATE: TX
;     COUNTRY: [illegible]
;     ZIP: 78701
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #[illegible] Version #[illegible]
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/[illegible]
;     FILING DATE: [illegible]
;     CLASSIFICATION: [illegible]
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: [illegible]
;     FILING DATE: [illegible] 1995
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Mayfield, Denise L.
;     REGISTRATION NUMBER: [illegible]
;     REFERENCE/DOCKET NUMBER: [illegible]
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: [illegible]
;     TELEFAX: 512/49[illegible]
; INFORMATION FOR SEQ ID NO: [illegible]:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: [illegible] amino acids
;     TYPE: amino acid
;     STRANDEDNESS: single
;     TOPOLOGY: linear
;   MOLECULE TYPE: peptide
US-08-659-254-1

  Query Match [illegible]
  Best Local Similarity [illegible]; Pred. No. [illegible]
  Matches [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible]

Qy          1 ALCILNN[illegible]WK 12
              |||||||||||
Db          1 ALCILNNQ[illegible]WK 12


RESULT 3
US-08-659-254-2
; Sequence 2, Application US/08659254
; Patent No. 6194[illegible]
; GENERAL INFORMATION:
;   APPLICANT: [illegible] M
;   APPLICANT: Clark, Dr. Barbara
;   TITLE OF INVENTION: COMPOSITIONS AND METHODS [illegible]
;   TITLE OF INVENTION: [illegible]
;   NUMBER OF SEQUENCES: [illegible]
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Ak[illegible]
;     STREET: 1900 Frost [illegible]
;     CITY: Austin
;     STATE: TX
;     COUNTRY: U.S.A.
```

```
;     [illegible]
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: [illegible] disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: [illegible]
;     SOFTWARE: [illegible] Release #[illegible] Version #[illegible]
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: [illegible]
;     FILING DATE: [illegible] JUN 1996
;     CLASSIFICATION: [illegible]
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: [illegible]
;     FILING DATE: [illegible]
;   ATTORNEY/AGENT INFORMATION:
;     NAME: [illegible]
;     REGISTRATION NUMBER: [illegible]
;     REFERENCE/DOCKET NUMBER: [illegible]
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: [illegible]
;     TELEFAX: [illegible]
; INFORMATION FOR SEQ ID NO: [illegible]
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: [illegible]
;     TYPE: [illegible]
;     STRANDEDNESS: [illegible]
;     TOPOLOGY: [illegible]
;   MOLECULE TYPE: [illegible]
US-08-659-254-[illegible]

  Query Match [illegible]
  Best Local Similarity [illegible]; Pred. No. [illegible]
  Matches [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible]

Qy          1 ALCILNNLFGWK [illegible]

Db          [illegible] ALCILNNQFGWK [illegible]


RESULT 4
[illegible]
; Sequence [illegible], Application US/08659254
; Patent No. [illegible]
; GENERAL INFORMATION:
;   APPLICANT: [illegible] M
;   APPLICANT: [illegible]
;   TITLE OF INVENTION: COMPOSITIONS AND METHODS [illegible]
;   TITLE OF INVENTION: [illegible]
;   NUMBER OF SEQUENCES: [illegible]
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Akin, Gump, Strauss, Hauer & Feld, L.L.P.
;     STREET: [illegible] Frost Bank [illegible] Avenue
;     CITY: Austin
;     STATE: TX
;     COUNTRY: [illegible]A
;     ZIP: [illegible]
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: [illegible] disk
;     COMPUTER: IBM PC [illegible]
;     OPERATING SYSTEM: [illegible]
;     SOFTWARE: [illegible] Version #[illegible]
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: [illegible]
;     FILING DATE: [illegible]
;     CLASSIFICATION: [illegible]
;   ATTORNEY/AGENT INFORMATION:
;     NAME: [illegible]
;     REGISTRATION NUMBER: [illegible]
;     REFERENCE/DOCKET NUMBER: [illegible]
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: [illegible]
;     TELEFAX: [illegible]
; INFORMATION FOR SEQ ID NO: [illegible]
;     SEQUENCE CHARACTERISTICS:
;        LENGTH:  284 amino acids
;        TYPE:  amino acid
;        STRANDEDNESS:  single
;        TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-538-960-2


  Query Match             100.0%;  Score 6[illegible];  DB 2;  Length 28[illegible];
  Best Local Similarity   100.0%;  Pred. No. [illegible].00021;
  Matches   12;  Conservative    0;  Mismatches    0;  Indels    [illegible];  Gaps    [illegible];

Qy        1 ALGILNNQEGWK 12
            ||||||||||||
Db       85 ALGILNNQEGWK 96


RESULT    5
US-08-410-540-2
; Sequence 2, Application US/08410540
; Patent No. 5807678
;  GENERAL INFORMATION:
;     APPLICANT:  Miller, Walter L.
;     APPLICANT:  Lin, Dong
;     APPLICANT:  Strauss III, Jerome F.
;     TITLE OF INVENTION:  IDENTIFICATION OF [illegible] GENE MUTATIONS
;     TITLE OF INVENTION.  ASSOCIATED WITH CONGENITAL LIPOID ADRENAL HYPERPLASIA
;     NUMBER OF SEQUENCES:  30
;     CORRESPONDENCE ADDRESS:
;        ADDRESSEE:  Cooley Godward Castro Huddleson & Tatum
;        STREET:  5 Palo Alto Square
;        CITY:  Palo Alto
;        STATE:  CA
;        COUNTRY:  US
;        ZIP:  94306-2155
;     COMPUTER READABLE FORM:
;        MEDIUM TYPE:  Floppy disk
;        COMPUTER:  IBM PC compatible
;        OPERATING SYSTEM:  PC DOS/MS DOS
;        SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;        APPLICATION NUMBER.  US/08/410,540
;        FILING DATE:  23-MAR-1995
;        CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;        NAME:  Neeley, Richard L.
;        REGISTRATION NUMBER:  30,092
;        REFERENCE/DOCKET NUMBER:  HCAL 2 [illegible]
;     TELECOMMUNICATION INFORMATION:
;        TELEPHONE:  415 853 5070
;        TELEFAX:  415 857 0663
;        TELEX:  380816COOLEYPA
;   INFORMATION FOR SEQ ID NO:  2:
;     SEQUENCE CHARACTERISTICS:
;        LENGTH:  285 amino acids
;        TYPE:  amino acid
;        TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-410-540-2


  Query Match             92.4%;  Score [illegible];  DB 1;  Length 285;
  Best Local Similarity   91.7%;  Pred. No. 0.0017;
  Matches   11;  Conservative    1;  Mismatches    0;  Indels    [illegible];  Gaps    [illegible];

Qy        1 ALGILNNQEGWK 12
            |||||:||||||
Db       86 ALGILSNQEGWK 97


RESULT    6
US-08-659-254-18
```

```
; Sequence 18, Application US/08659254
; Patent No. 6194555
;  GENERAL INFORMATION:
;     APPLICANT:  Stocco, Douglas M.
;     APPLICANT:  Clark, Dr. Barbara J.
;     TITLE OF INVENTION.  COMPOSITIONS AND METHODS FOR REGULATION
;     TITLE OF INVENTION.  OF STEROIDOGENESIS
;     NUMBER OF SEQUENCES:  19
;     CORRESPONDENCE ADDRESS:
;        ADDRESSEE:  Akin, Gump, Strauss, Hauer & Feld, L.L.P.
;        STREET:  1900 Frost Bank Plaza, 816 Congress Avenue
;        CITY:  Austin
;        STATE:  TX
;        COUNTRY.  U.S.A.
;        ZIP:  78701
;     COMPUTER READABLE FORM:
;        MEDIUM TYPE:  Floppy disk
;        COMPUTER:  IBM PC compatible
;        OPERATING SYSTEM.  PC DOS/MS-DOS
;        SOFTWARE:  PatentIn Release #1.0, Version #1.[illegible]
;     CURRENT APPLICATION DATA:
;        APPLICATION NUMBER.  US/08/659,254
;        FILING DATE:  07-JUN-1996
;        CLASSIFICATION:  530
;     PRIOR APPLICATION DATA:
;        APPLICATION NUMBER:  US 08/538,960
;        FILING DATE   04-OCT-1995
;     ATTORNEY/AGENT INFORMATION:
;        NAME:  Mayfield, Denise L.
;        REGISTRATION NUMBER:  33,732
;        REFERENCE/DOCKET NUMBER:  43-75 [illegible] 6
;     TELECOMMUNICATION INFORMATION:
;        TELEPHONE:  512/499-6200
;        TELEFAX:  512/499-6290
;   INFORMATION FOR SEQ ID NO:  18:
;     SEQUENCE CHARACTERISTICS:
;        LENGTH:  285 amino acids
;        TYPE:  amino acid
;        TOPOLOGY:  linear
US-08-659-254-18


  Query Match             92.4%;  Score [illegible];  DB 4;  Length 285;
  Best Local Similarity   91.7%;  Pred. No. [illegible]17;
  Matches   11;  Conservative    1;  Mismatches    0;  Indels    [illegible];  Gaps    0;

Qy        1 ALGILNNQEGWK 12
            |||||:||||||
Db       86 ALGILSNQEGWK 97


RESULT    7
US-08-538-960-8
; Sequence 8, Application US/08538960
; Patent No. 5872230
;  GENERAL INFORMATION:
;     APPLICANT:  Stocco, Douglas M.
;     APPLICANT:  Clark, Barbara J.
;     TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR
;     TITLE OF INVENTION:  REGULATION OF STEROIDOGENESIS
;     NUMBER OF SEQUENCES:  14
;     CORRESPONDENCE ADDRESS:
;        ADDRESSEE:  Akin, Gump, Strauss, Hauer & Feld, L.L.P.
;        STREET:  1900 Frost Bank Plaza, 816 Congress Avenue
;        CITY:  Austin
;        STATE:  TX
;        COUNTRY:  U.S.A.
;        ZIP:  78701
;     COMPUTER READABLE FORM:
;        MEDIUM TYPE:  Floppy disk
;        COMPUTER:  IBM PC compatible
;        OPERATING SYSTEM.  PC DOS/MS-DOS
;        SOFTWARE:  PatentIn Release #1.0, Version #1.[illegible]
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER: US/08/[illegible]
;      FILING DATE: [illegible]
;      CLASSIFICATION: [illegible]
;    ATTORNEY/AGENT INFORMATION:
;      NAME: Mayfield, [illegible]
;      REGISTRATION NUMBER: [illegible]
;      REFERENCE/DOCKET NUMBER: [illegible]
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE: [illegible]
;      TELEFAX: [illegible]
;  INFORMATION FOR SEQ ID NO: 8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH: 11 amino acids
;      TYPE: amino acid
;      STRANDEDNESS: [illegible]
;      TOPOLOGY: linear
;    MOLECULE TYPE: peptide
[illegible]538 [illegible] 8

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches [illegible]

Qy          4 [illegible]NNQEWK [illegible]
              [illegible]
Db          1 [illegible]NNQEWK [illegible]


RESULT 8
US-08-659-254-8
; Sequence 8, Application US/08659254
; Patent No. 6194[illegible]
; GENERAL INFORMATION:
;    APPLICANT: [illegible] M
;    APPLICANT: Clark, [illegible] Barbara J
;    TITLE OF INVENTION: COMPOSITIONS AND METHODS [illegible]
;    TITLE OF INVENTION: OF STEROIDOGENESIS
;    NUMBER OF SEQUENCES: [illegible]
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE: [illegible]
;      STREET: [illegible] Frost Bank [illegible]
;      CITY: Austin
;      STATE: TX
;      COUNTRY: [illegible]
;      ZIP: 78701
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE: Floppy disk
;      COMPUTER: [illegible]
;      OPERATING SYSTEM: [illegible]/MS-[illegible]
;      SOFTWARE: [illegible]
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER: [illegible]
;      FILING DATE: [illegible]
;      CLASSIFICATION: [illegible]
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER: [illegible]
;      FILING DATE: [illegible]
;    ATTORNEY/AGENT INFORMATION:
;      NAME: Mayfield, [illegible]
;      REGISTRATION NUMBER: [illegible]
;      REFERENCE/DOCKET NUMBER: [illegible]
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE: [illegible]
;      TELEFAX: [illegible]
;  INFORMATION FOR SEQ ID NO: [illegible]
;    SEQUENCE CHARACTERISTICS:
;      LENGTH: 11 [illegible]
;      TYPE: amino acid
;      STRANDEDNESS: [illegible]
;      TOPOLOGY: linear
;    MOLECULE TYPE: [illegible]
```

```
US-[illegible]

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches [illegible] Mismatches [illegible] Indels [illegible] Gaps [illegible]

Qy          4 [illegible]NNQEWK [illegible]

Db          1 [illegible]NNQEWK [illegible]


RESULT [illegible]
US-[illegible]
; Sequence [illegible], Application [illegible]
; Patent No. [illegible]
; GENERAL INFORMATION:
;    APPLICANT: [illegible] M
;    APPLICANT: [illegible] Barbara [illegible]
;    TITLE OF INVENTION: [illegible] AND METHODS [illegible]
;    TITLE OF INVENTION: [illegible]
;    NUMBER OF SEQUENCES: [illegible]
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE: [illegible]
;      STREET: [illegible]
;      CITY: [illegible]
;      STATE: TX
;      COUNTRY: [illegible]
;      ZIP: [illegible]
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE: [illegible]
;      COMPUTER: [illegible]
;      OPERATING SYSTEM: [illegible]
;      SOFTWARE: [illegible]
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER: [illegible]
;      FILING DATE: Concurrently herewith
;      CLASSIFICATION: [illegible]
;    ATTORNEY/AGENT INFORMATION:
;      NAME: [illegible]
;      REGISTRATION NUMBER: [illegible]
;      REFERENCE/DOCKET NUMBER: [illegible]
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE: [illegible]
;      TELEFAX: [illegible]
;  INFORMATION FOR SEQ ID NO: [illegible]
;    SEQUENCE CHARACTERISTICS:
;      LENGTH: [illegible]
;      TYPE: [illegible]
;      STRANDEDNESS: [illegible]
;      TOPOLOGY: [illegible]
;    MOLECULE TYPE: peptide
[illegible]

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches [illegible] Mismatches [illegible] Indels [illegible] Gaps [illegible]

Qy            [illegible]NQEWK [illegible]

Db            [illegible]NQEWK [illegible]


RESULT [illegible]
US-[illegible]
; Sequence [illegible]
; Patent No. [illegible]
; GENERAL INFORMATION:
;    APPLICANT: [illegible] M
;    APPLICANT: [illegible]
;    TITLE OF INVENTION: [illegible] AND METHODS FOR REGULATION
;    TITLE OF INVENTION: [illegible]
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Akin, Gump, Strauss, Hauer & Feld, L.L.P.
;      STREET:  1900 Frost Bank Plaza, 816 Congress Avenue
;      CITY:  Austin
;      STATE:  TX
;      COUNTRY:  U.S.A.
;      ZIP:  78701
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/659,254
;      FILING DATE:  07-JUN-1996
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/538,96[illegible]
;      FILING DATE:  04-OCT-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Mayfield, Denise L.
;      REGISTRATION NUMBER:  33,732
;      REFERENCE/DOCKET NUMBER:  43375 00[illegible]
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  512/499-6200
;      TELEFAX:  512/499-6290
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  6 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-659-254-6

  Query Match             57.6%;  Score 38;  DB 4;  Length 6;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+05;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          7 NQEGWK 12
              ||||||
Db          1 NQEGWK 6


RESULT 11
US-09-026-587-3
; Sequence 3, Application US/09026587
; Patent No. 5912128
;  GENERAL INFORMATION:
;    APPLICANT:  Lal, Preeti
;    APPLICANT:  Guegler, Karl J.
;    APPLICANT:  Corley, Neil C.
;    TITLE OF INVENTION:  HUMAN ENA/VASP-LIKE PROTEIN
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/026,587
;      FILING DATE:  Filed Herewith
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-047[illegible] US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-855-0555
;      TELEFAX:  650-845-4166
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  393 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  GenBank
;      CLONE:  164445[illegible]
US-09-026-587-3

  Query Match             56.1%;  Score 37;  DB 2;  Length 393;
  Best Local Similarity   70.0%;  Pred. No. 4[illegible];
  Matches    7;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 ALCIENNQEG 10
              || |:|:|||
Db        108 ALNIMNSQEG 117


RESULT 12
US-09-227-420-3
; Sequence 3, Application US/09227420
; Patent No. 5990087
;  GENERAL INFORMATION:
;    APPLICANT:  Lal, Preeti
;    APPLICANT:  Guegler, Karl J.
;    APPLICANT:  Corley, Neil C.
;    TITLE OF INVENTION:  HUMAN ENA/VASP-LIKE PROTEIN
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/227,420
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  09/026,587
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-047[illegible] US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-855-0555
;      TELEFAX:  650-845-4166
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  393 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;          TOPOLOGY: linear
;       IMMEDIATE SOURCE:
;          LIBRARY: Genbank
;          CLONE: 164414
US-09-227-426-3

  Query Match         [illegible]
  Best Local Similarity [illegible]
  Matches    7;  Conservative    2;  Mismatches [illegible]

Qy          1 ALTLNNQ[illegible]
                || ||||||
Db        108 ALNTMNSQ[illegible]


RESULT 13
[illegible] 026 587 1
; Sequence 1, Application US/09026587
; Patent No. 5912128
;  GENERAL INFORMATION:
;     APPLICANT: Lal, Preeti
;     APPLICANT: Guegler, Karl J.
;     APPLICANT: Corley, Neil C.
;     TITLE OF INVENTION: HUMAN ENA/VASP-LIKE PROTEIN
;     NUMBER OF SEQUENCES: 4
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE: Incyte Pharmaceuticals, Inc.
;       STREET: 3174 Porter Drive
;       CITY: Palo Alto
;       STATE: CA
;       COUNTRY: USA
;       ZIP: 94304
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE: Diskette
;       COMPUTER: IBM Compatible
;       OPERATING SYSTEM: DOS
;       SOFTWARE: FastSEQ for Windows Version [illegible]
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER: US/09/026,587
;       FILING DATE: Filed Herewith
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:
;       FILING DATE:
;     ATTORNEY/AGENT INFORMATION:
;       NAME: Billings, Lucy J.
;       REGISTRATION NUMBER: [illegible]
;       REFERENCE/DOCKET NUMBER: PF-[illegible]
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE: 650-855-0555
;       TELEFAX: 650-845-4166
;       TELEX:
;  INFORMATION FOR SEQ ID NO: 1:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH: 418 amino acids
;       TYPE: amino acid
;       STRANDEDNESS: single
;       TOPOLOGY: linear
;     IMMEDIATE SOURCE:
;       LIBRARY: BRAINOT03
;       CLONE: 308411
US-09-026-587-1

  Query Match         [illegible]
  Best Local Similarity [illegible]
  Matches    7;  Conservative [illegible]  Mismatches [illegible]

Qy          1 ALTLNNQ[illegible]
                |   |||
Db        110 ALNTMNSQ[illegible]
```

```
RESULT 14
US-09-227-426-4
; Sequence 4, Application US/09227426
; Patent No. [illegible]
;  GENERAL INFORMATION:
;     APPLICANT: Lal, Preeti
;     APPLICANT: Guegler, Karl J.
;     APPLICANT: Corley, Neil C.
;     TITLE OF INVENTION: HUMAN ENA/VASP-LIKE PROTEIN
;     NUMBER OF SEQUENCES: 4
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE: Incyte Pharmaceuticals, Inc.
;       STREET: 3174 Porter Drive
;       CITY: Palo Alto
;       STATE: CA
;       COUNTRY: USA
;       ZIP: 94304
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE: Diskette
;       COMPUTER: IBM Compatible
;       OPERATING SYSTEM: DOS
;       SOFTWARE: FastSEQ for Windows Version 2.0
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER: US/09/227,426
;       FILING DATE:
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER: 09/026,587
;       FILING DATE:
;     ATTORNEY/AGENT INFORMATION:
;       NAME: Billings, Lucy J.
;       REGISTRATION NUMBER: [illegible]
;       REFERENCE/DOCKET NUMBER: [illegible]
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE: [illegible]
;       TELEFAX: [illegible]
;       TELEX:
;  INFORMATION FOR SEQ ID NO: 4:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH: 418 amino acids
;       TYPE: amino acid
;       STRANDEDNESS: single
;       TOPOLOGY: linear
;     IMMEDIATE SOURCE:
;       LIBRARY: BRAINOT03
;       CLONE: 308411
US-09-227-426-4

  Query Match         [illegible]  Length 418;
  Best Local Similarity [illegible]  Pred. No. [illegible]
  Matches [illegible]  Conservative [illegible]  Mismatches [illegible]  Indels    0;  Gaps    0;

Qy          1 ALTLNNQ[illegible]
                |   |||
Db          [illegible] ALNTMNSQ[illegible]


RESULT 15
[illegible]
; Sequence [illegible], Application [illegible]
; Patent No. [illegible]
;  GENERAL INFORMATION:
;     APPLICANT: WALTON, JAMES [illegible]
;     APPLICANT: [illegible]
;     TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE TREATMENT OF IMMUNOLOGICALLY
;     NUMBER OF SEQUENCES: [illegible]
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE: [illegible]
;       STREET: [illegible]
;       CITY: [illegible]
;       STATE: [illegible]
;       COUNTRY:  USA
;       ZIP:  98121
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Diskette
;       COMPUTER:  IBM Compatible
;       OPERATING SYSTEM:  DOS
;       SOFTWARE:  FastSEQ for Windows Version 2.0
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/997,080
;       FILING DATE:
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:
;       FILING DATE:
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Sleath, Janet
;       REGISTRATION NUMBER:  37,007
;       REFERENCE/DOCKET NUMBER:  11000.10[illegible]
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  206-269-0565
;       TELEFAX:  206-269-0563
;       TELEX:
;   INFORMATION FOR SEQ ID NO:  158:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  161 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
US-08-997-080-158

  Query Match             54.5%;  Score 30;  DB 2;  Length 161;
  Best Local Similarity   66.7%;  Pred. No. 25;
  Matches    6;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          3 GILNNQEGW 11
              | ||  |||
Db        103 GFLNPSEGW 111


Search completed: August 22, 2001, 14:30:[illegible]
Job time: 364 sec
```

```
                         GenCore version 4.5
              Copyright (c) 1993 - 2001 Compugen Ltd.

OM protein - protein search, using sw model

Run on:          August 22, 2001, 14:30:52   Search time 35.8 Seconds
                                              (without alignments)
                                              [illegible] Million cell updates/sec

Title:           US-08-659-254-3
Perfect score:   80
Sequence:        1 AEHGPTCMVLHPLA 14

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters:      93435

   mum DB seq length: 0
   mum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       SwissProt_39:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                              SUMMARIES
                %
Result          Query
  No.   Score   Match  Length DB  ID                          Description
-------------------------------------------------------------------------------
   1      80    100.0    284   1  STAR_MOUSE                  P51557 mus musculu
   2      80    100.0    284   1  STAR_RAT                    P97826 rattus norv
   3      80    100.0    285   1  STAR_HORSE                  O46689 equus cabal
   4      80    100.0    285   1  STAR_HUMAN                  P49675 homo sapien
   5      80    100.0    285   1  STAR_PIG                    Q28996 sus scrofa
   6      72     90.0    284   1  STAR_MESAU                  P79114 mesocricetu
   7      72     90.0    285   1  STAR_BOVIN                  Q28918 bos taurus
   8    43.5     54.4    323   1  PPA5_HUMAN                  P13686 homo sapien
   9    43.5     54.4    327   1  PPA5_MOUSE                  Q05117 mus musculu
  10    43.5     54.4    327   1  PPA5_RAT                    P29288 rattus norv
  11      43     53.8    670   1  IKTP_ALC[illegible]         P21726 alcaligenes
  12      42     52.5    549   1  LIP4_CANRU                  P32948 candida rug
  13      41     51.2    246   1  NEUM_CHICK                  [illegible] gallus gall
  14    40.5     50.6    340   1  PPA5_PIG                    P09889 sus scrofa
  15      40     50.0    479   1  PRL2_ARATH                  Q39[illegible] arabidopsis
  16      40     50.0    548   1  LIP2_CANRU                  P32946 candida rug
  17      40     50.0   4543   1  LRP1_CHICK                  P98157 gallus gall
  18      38     47.5    193   1  RECR_HELPJ                  Q9zks6 helicobacte
  19      38     47.5    193   1  RECR_HELPY                  P56211 helicobacte
  20      38     47.5    377   1  NUEM_HUMAN                  Q16795 homo sapien
  21      38     47.5   1286   1  PATC_DROME                  P18502 drosophila
  22      37     46.2    204   1  YACE_A[illegible]           [illegible]
  23      37     46.2    275   1  NIRQ_PSE[illegible]         Q02441 pseudomonas
  24      37     46.2    395   1  YI44_STR[illegible]         P20188 streptomyce
  25      37     46.2    426   1  POLG_CRPV                   P13418 cricket par
  26      37     46.2    572   1  HEMO_MOUSE                  [illegible] mus musculu
  27      37     46.2    587   1  HEMO_RAT                    Q63147 rattus norv
  28      37     46.2    670   1  IKIC_AL[illegible]          P21725 alcaligenes
  29      37     46.2    716   1  HGFL_MOUSE                  P26928 mus musculu
  30      37     46.2   1075   1  NFC3_MOUSE                  P97[illegible] mus musculu
  31      36     45.0    162   1  YB09_SCHPO                  P87150 schizosacch
  32      36     45.0    229   1  LIPB_XYLFA                  Q9PD[illegible] xylella fas
  33      36     45.0    304   1  VIRA_VIBAN                  Q56[illegible] vibrio ang
  34      36     45.0    357   1  AAAA_EMENI                  P21133 emericella
  35      36     45.0    474   1  A1BG_HUMAN                  P04217 homo sapien
  36      36     45.0    614   1  DDX5_HUMAN                  P17844 homo sapien
  37      36     45.0    614   1  DDX5_MOUSE                  Q61656 mus musculu
  38      36     45.0    622   1  RRP2_THOGV                  P27[illegible] thogoto vir
  39      36     45.0    650   1  DD17_HUMAN                  Q92841 homo sapien
  40    35.5     44.4    490   1  Y888_MYCTU                  [illegible] mycobacteri
  41      35     43.8    168   1  AIPL_RAT                    P35434 rattus norv
  42      35     43.8    242   1  CRP1_LUMRU                  P06[illegible] lumbricus pol
  43      35     43.8    242   1  CRP[illegible]              [illegible] lumbricus pol
  44      35     43.8    374   1  FOS_TET[illegible]          [illegible] tetrahymen
  45      35     43.8    462   1  YNEI_ECOLI                  P76149 escherichia

                              ALIGNMENTS

RESULT 1
STAR_MOUSE
ID   STAR_MOUSE     STANDARD;      PRT;   284 AA.
AC   P51557;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN PRECURSOR (LUTEINIZING HORMONE
DE   INDUCED PROTEIN).
GN   STAR.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID 10090;
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL [illegible]
RX   MEDLINE 950509[illegible]6, PubMed 7961770;
RA   Clark B.J., Wells J., King S.R., Stocco D.M.;
RT   "The purification, cloning, and expression of a novel luteinizing
RT   hormone-induced mitochondrial protein in MA-10 mouse Leydig tumor
RT   cells. Characterization of the steroidogenic acute regulatory protein
RT   (StAR).";
RL   J. Biol. Chem. 269:28314-28322(1994).
CC   -!- FUNCTION: PLAYS A KEY ROLE IN STEROID HORMONE SYNTHESIS BY
CC       ENHANCING THE METABOLISM OF CHOLESTEROL INTO PREGNENOLONE.
CC       TRANSPORTER THAT BINDS TO AND TRANSPORT CHOLESTEROL THROUGH THE
CC       INTERMEMBRANE SPACE OF THE MITOCHONDRION (PROBABLE).
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL.
CC   -!- SIMILARITY: BELONGS TO THE STAR FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; L36062; AAB94783.1; -.
DR   MGD; MGI:102760; Star.
DR   InterPro; IPR000799; -.
DR   InterPro; IPR002913; -.
DR   Pfam; PF01852; START; 1.
DR   PRINTS; PR00978; STARPROTEIN.
KW   Lipid-binding; Lipid transport; Transport; Steroidogenesis;
KW   Mitochondrion; Transit peptide.
FT   TRANSIT       [illegible]   ?       MITOCHONDRION (POTENTIAL).
FT   CHAIN         [illegible]   284     STEROIDOGENIC ACUTE REGULATORY PROTEIN.
FT   DOMAIN        [illegible]   174     [illegible]
SQ   SEQUENCE   284 AA;  31626 MW;  [illegible] CRC64;

  Query Match             100.0%;  Score 80;  DB 1;  Length 284;
  Best Local Similarity   100.0%;  Pred. No. [illegible];
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          1 AEHGPIMVIPLA 1[illegible]
              ||||||||||||
Db        217 AEHGPIMVIPLA 2[illegible]

RESULT 2
STAR_RAT
ID   STAR_RAT       STANDARD;      PRT;   284 AA.
AC   P97826; P97[illegible]
DT   01-NOV-1997 (Rel. [illegible], Created)
DT   01-NOV-1997 (Rel. [illegible], Last sequence update)
DT   30-MAY-2000 (Rel. [illegible], Last annotation update)
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN, [illegible]
GN   STAR.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; [illegible]
OC   Mammalia; Eutheria; Rodentia; [illegible] Muridae; [illegible]
OX   NCBI_TaxID=[illegible];
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=WISTAR; TISSUE=[illegible];
RX   MEDLINE=974[illegible]; PubMed=[illegible];
RA   Mizutani T., Sakata [illegible], Minegishi [illegible], [illegible]
RT   "Molecular cloning, characterization and cellular [illegible]
RT   steroidogenic acute regulatory protein (StAR) [illegible]
RL   Life Sci. 61:14[illegible](1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY; TISSUE=Testis;
RX   MEDLINE=9716[illegible]; PubMed=[illegible];
RA   Lee H.K., Ahn R.[illegible], [illegible]
RT   "Nucleotide sequence of rat steroidogenic acute regulatory protein
RT   complementary DNA.";
RL   Biochem. Biophys. Res. Commun. [illegible]
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY; TISSUE=Adrenal [illegible];
RX   MEDLINE=9818[illegible]; PubMed=[illegible];
RA   Ariyoshi N., Kim Y.[illegible], Artemenko [illegible]
RA   Jefcoate C.R.;
RT   "Characterization of the rat Star gene [illegible]
RT   3.5 kilobase pair mRNA: ACTH [illegible]
RT   precedes elevation of Star mRNA and protein.";
RL   J. Biol. Chem. [illegible]-7619(1998).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY; TISSUE=Adrenal [illegible];
RX   MEDLINE=9904[illegible]; PubMed=[illegible];
RA   Furukawa A., Miyatake A., [illegible]
RT   "Steroidogenic acute regulatory protein (StAR) [illegible]
RT   constitutively expressed in the adult rat [illegible]
RT   colocalization of StAR, cytochrome P-450scc (CYPXIA1) and [illegible]
RT   hydroxysteroid dehydrogenase in the rat brain.";
RL   J. Neurochem. 71:2231-2238(1998).
CC   -!- FUNCTION: PLAYS A KEY ROLE IN STEROID HORMONE SYNTHESIS BY
CC       ENHANCING THE METABOLISM OF CHOLESTEROL INTO PREGNENOLONE.
CC       TRANSPORTER THAT [illegible] AND TRANSFER [illegible]
CC       INTERMEMBRANE SPACE OF THE MITOCHONDRION (PROBABLE).
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL.
CC   -!- SIMILARITY: BELONGS TO THE STAR FAMILY.
CC   --------
CC   This SWISS-PROT entry is copyright. [illegible]
CC   between the Swiss Institute [illegible]
CC   the European Bioinformatics Institute. There are [illegible]
CC   use by non-profit [illegible]
CC   modified and this statement is not removed. [illegible]
CC   entities requires a license agreement [illegible]
CC   or send an email to [illegible]
CC   --------
DR   EMBL; AB001349; BAA19[illegible]
DR   EMBL; U76419; AAB1[illegible]
DR   EMBL; AF044081; AAC0[illegible]
DR   EMBL; AB00608[illegible]; BAA34[illegible]
```

```
DR   [illegible]
DR   InterPro; [illegible]
DR   Pfam; [illegible] START; [illegible]
DR   PRINTS; PR00978; STARPROTEIN.
KW   [illegible] Transport; [illegible] Steroidogenesis;
KW   Mitochondrion; Transit peptide.
FT   TRANSIT      [illegible]     MITOCHONDRION (POTENTIAL).
FT   CHAIN        [illegible] 284 STEROIDOGENIC ACUTE REGULATORY PROTEIN.
FT   [illegible]                  M[illegible]LATEK [illegible] AV[illegible]YV[illegible] [illegible]
FT                                [illegible]VYPG[illegible] [illegible] VVVNPYIAP[illegible]
FT                                [illegible]
SQ   SEQUENCE   284 AA;  [illegible] MW;  [illegible] CRC64;

  Query Match  [illegible]  Score [illegible]  DB [illegible]  Length 284;
  Best Local Similarity [illegible]  Pred. No. [illegible]
  Matches [illegible]  Conservative [illegible]  Mismatches [illegible]  Indels [illegible]  Gaps [illegible]

Qy          [illegible] AEHGPIMVIPLA [illegible]
              [illegible]
Db        [illegible] AEHGPIMVIPLA [illegible]


RESULT 3
STAR_HORSE
ID   STAR_HORSE     STANDARD;      PRT;   [illegible] AA.
AC   [illegible]
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   30-MAY-2000 (Rel. [illegible], Last annotation update)
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN PRECURSOR.
GN   STAR.
OS   Equus caballus (Horse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX   NCBI_TaxID=[illegible];
RN   [1]
RP   SEQUENCE FROM N.A.
RC   [illegible]
RX   MEDLINE=[illegible]; PubMed=[illegible];
RA   Kerban A., [illegible]
RT   "Human chorionic gonadotropin induces an inverse regulation of
RT   steroidogenic acute regulatory protein messenger ribonucleic acid in
RT   theca interna and granulosa cells of equine preovulatory follicles.";
RL   Endocrinology [illegible](1999).
CC   -!- FUNCTION: PLAYS A KEY ROLE IN STEROID HORMONE SYNTHESIS BY
CC       ENHANCING THE METABOLISM OF CHOLESTEROL INTO PREGNENOLONE.
CC       TRANSPORTER THAT BINDS [illegible] AND TRANSFER [illegible] THROUGH THE
CC       INTERMEMBRANE SPACE OF THE MITOCHONDRION (PROBABLE).
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL.
CC   -!- SIMILARITY: BELONGS TO THE STAR FAMILY.
CC   --------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. [illegible]
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to [illegible]
CC   --------
DR   EMBL; AF[illegible]; AAC[illegible]
DR   EMBL; AF[illegible]; AAD[illegible]
DR   [illegible]
DR   InterPro; [illegible]
DR   Pfam; [illegible] START; [illegible]
DR   PRINTS; PR00978; STARPROTEIN.
KW   [illegible] Transport; [illegible] Steroidogenesis;
KW   Mitochondrion; Transit peptide.
FT   TRANSIT      [illegible]     MITOCHONDRION (POTENTIAL).
FT   CHAIN        [illegible]     STEROIDOGENIC ACUTE REGULATORY PROTEIN.
SQ   SEQUENCE   [illegible] AA;  [illegible] MW;  [illegible] CRC64;
```

```
  Query Match             100.0%;  Score [illegible];  DB 1;  Length 285;
  Best Local Similarity   100.0%;  Pred. No. [illegible]-06;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AEHGPTCMVLHPLA 14
              ||||||||||||||
Db        218 AEHGPTCMVLHPLA 231


RESULT   4
STAR_HUMAN
ID   STAR_HUMAN     STANDARD;      PRT;   285 AA.
AC   P49675; Q16396;
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN PRECURSOR.
GN   STAR.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
     [1]
     SEQUENCE FROM N.A.
RC   TISSUE=Adrenal cortex;
RX   MEDLINE=95281540; PubMed=7761400;
RA   Sugawara T., Holt J.A., Driscoll D., Strauss J.F. III, Lin D.,
RA   Miller W.L., Patterson D., Clancy K.P., Hart I.M., Clark B.J.,
RA   Stocco D.M.;
RT   "Human steroidogenic acute regulatory protein: functional activity in
RT   COS-1 cells, tissue-specific expression, and mapping of the
RT   structural gene to 8p11.2 and a pseudogene to chromosome 13.";
RL   Proc. Natl. Acad. Sci. U.S.A. 92:4778-4782(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Placenta;
RX   MEDLINE=96038208; PubMed=7547998;
RA   Sugawara T., Lin D., Holt J.A., Martin K.O., Javitt N.B., Miller W.L.,
RA   Strauss J.F. III;
RT   "Structure of the human steroidogenic acute regulatory protein (StAR)
RT   gene: StAR stimulates mitochondrial cholesterol 27-hydroxylase
RT   activity.";
RL   Biochemistry 34:12506-12512(1995).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96008588; PubMed=7548191;
RA   Gradi A., Tang-Wai R., McBride H.M., Chu L.L., Shore G.C.,
RA   Pelletier J.;
RT   "The human steroidogenic acute regulatory (StAR) gene is expressed in
RT   the urogenital system and encodes a mitochondrial polypeptide.";
RL   Biochim. Biophys. Acta 1258:228-233(1995).
     [4]
     SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RA   Yu W., Sarginson J., Gibbs R.A.;
RL   Submitted (NOV-1997) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   VARIANT CLAH LEU-182.
RX   MEDLINE=97096837; PubMed=8948562;
RA   Bose H.S., Sugawara T., Strauss J.F. III, Miller W.L.;
RT   "The pathophysiology and genetics of congenital lipoid adrenal
RT   hyperplasia.";
RL   New Engl. J. Med. 335:1870-1878(1996).
RN   [6]
RP   VARIANT ALA-203.
RX   MEDLINE=98112495; PubMed=9452116;
RA   Katsumata N., Tanae A., Shinagawa T., Nagashima-Miyokawa A.,
RA   Shimizu M., Yasunaga T., Tanaka T., Hibi I.;
RT   "A novel frameshift mutation 840delA and a novel polymorphism [illegible] in
RT   the steroidogenic acute regulatory protein gene in a Japanese patient
RT   with congenital lipoid adrenal hyperplasia.";
RL   Hum. Mutat. Suppl. 1:S304-S307(1998).
CC   -!- FUNCTION: PLAYS A KEY ROLE IN STEROID HORMONE SYNTHESIS BY
CC       ENHANCING THE METABOLISM OF CHOLESTEROL INTO PREGNENOLONE.
CC       TRANSPORTER THAT BINDS TO AND TRANSFERS CHOLESTEROL THROUGH THE
CC       INTERMEMBRANE SPACE OF THE MITOCHONDRION (PROBABLE).
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN [illegible] ADRENAL [illegible] AND
CC       KIDNEY.
CC   -!- DISEASE: DEFECTS IN STAR CAUSE CONGENITAL LIPOID ADRENAL
CC       HYPERPLASIA (CLAH), THE MOST SEVERE FORM OF ADRENAL HYPERPLASIA.
CC   -!- SIMILARITY: BELONGS TO THE STAR FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are  no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its  content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; U17280; AAC50141.1; -.
DR   EMBL; U29105; AAC50234.1; -.
DR   EMBL; U29099; AAC50234.1; JOINED.
DR   EMBL; U29100; AAC50234.1; JOINED.
DR   EMBL; U29101; AAC50234.1; JOINED.
DR   EMBL; U29102; AAC50234.1; JOINED.
DR   EMBL; U29103; AAC50234.1; JOINED.
DR   EMBL; U29104; AAC50234.1; JOINED.
DR   EMBL; S79669; AAB35726.1; -.
DR   EMBL; AF03[illegible]; AAB88174.1; -.
DR   MIM; 600617; -.
DR   MIM; 201710; -.
DR   InterPro; IPR000799; -.
DR   InterPro; IPR002913; -.
DR   Pfam; PF01852; START; 1.
DR   PRINTS; PR00978; STARPROTEIN.
KW   Lipid binding; Lipid transport; Transport; Steroidogenesis;
KW   Mitochondrion; Transit peptide; Disease mutation; Polymorphism.
FT   TRANSIT       1     ?       MITOCHONDRION (POTENTIAL).
FT   CHAIN         ?   285       STEROIDOGENIC ACUTE REGULATORY PROTEIN.
FT   VARIANT     182   182       R -> L (IN CLAH).
FT                               /FTId=VAR_[illegible].
FT   VARIANT     203   203       D -> A.
FT                               /FTId=VAR_[illegible].
SQ   SEQUENCE   285 AA;  31953 MW;  [illegible] CRC64;

  Query Match             100.0%;  Score [illegible];  DB 1;  Length 285;
  Best Local Similarity   100.0%;  Pred. No. [illegible]-06;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AEHGPTCMVLHPLA 14
              |||||||| |||||
Db        218 AEHGPTCMVLHPLA 231


RESULT   5
STAR_PIG
ID   STAR_PIG       STANDARD;      PRT;   285 AA.
AC   Q28996; Q95254;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN PRECURSOR.
GN   STAR.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Ovary;
RX   MEDLINE=[illegible]; PubMed=[illegible];
RA   Pilon N., Daneau I., Brisson C., [illegible], Lussier J.G.,
RA   Silversides D.W.;
```

```
RT   "Porcine and bovine steroidogenic [illegible]
RT   gene expression during [illegible]";
RL   Endocrinology [illegible](1997).
RN   [2]
RP   SEQUENCE OF [illegible]-210 FROM N.A.
RC   TISSUE=Corpus luteum;
RX   MEDLINE=97[illegible]; PubMed=[illegible];
RA   Balasubramanian K., [illegible] H.A., [illegible]
RA   Veldhuis J.D.;
RT   "Regulation of porcine granulosa [illegible] steroidogenic [illegible]
RT   protein (StAR) [illegible]
RT   follicle stimulating hormone [illegible] protein kinase A [illegible]";
RL   Endocrinology [illegible]
CC   -!- FUNCTION: PLAYS A KEY ROLE IN STEROID HORMONE [illegible]
CC       ENHANCING THE METABOLISM OF CHOLESTEROL INTO PREGNENOLONE.
CC       TRANSPORTER THAT BINDS [illegible] AND TRANSFERS [illegible]
CC       INTERMEMBRANE SPACE OF THE MITOCHONDRION (PROBABLE).
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN ADULT AND FETAL [illegible]
CC       TISSUES, INCLUDING ADULT TESTES AND OVARIES AND ADULT ADRENAL
CC       GLANDS AS WELL AS STEROIDOGENIC [illegible]
CC       DEVELOPING FETAL TESTES, [illegible]
CC       THE FETAL OVARY.
CC   -!- SIMILARITY: BELONGS TO THE STAR FAMILY.
CC   -
CC   This SWISS-PROT entry is copyright. [illegible]
CC   between the Swiss Institute of Bioinformatics [illegible]
CC   the European Bioinformatics Institute. [illegible]
CC   use by non-profit institutions [illegible] way
CC   modified and this statement is not removed. [illegible]
CC   entities requires a license agreement (See http://www.[illegible]/announce/
CC   or send an email to [illegible]).
CC
DR   EMBL; [illegible]; AAC[illegible]
DR   EMBL; [illegible]195; AAB[illegible]
DR   InterPro; IPR[illegible];
DR   InterPro; IPR00[illegible];
DR   Pfam; PF0185[illegible]; START; [illegible]
DR   PRINTS; PR00[illegible]8; STARPROTEIN.
KW   Lipid-binding; Lipid transport; Transport; [illegible]
KW   Mitochondrion; Transit peptide.
FT   TRANSIT       1     [illegible]       MITOCHONDRION (POTENTIAL).
FT   CHAIN         [illegible]    28[illegible]       STEROIDOGENIC ACUTE REGULATORY PROTEIN.
FT   CONFLICT    198    198       [illegible]
SQ   SEQUENCE   28[illegible] AA;  [illegible] MW;  [illegible]

  Query Match          [illegible]
  Best Local Similarity   [illegible];  Pred. No. [illegible];
  Matches   [illegible];  Conservative   [illegible];  Mismatches   [illegible]

Qy          1 AEHGPICMVLH[illegible]A [illegible]
              |||||||||
Db        218 AEHGPICMVLH[illegible]A [illegible]


RESULT 6
STAR_MESAU
ID   STAR_MESAU     STANDARD;      PRT;   28[illegible] AA.
AC   P70114;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN [illegible]
GN   STAR.
OS   Mesocricetus auratus (Golden hamster).
OC   Eukaryota; Metazoa; [illegible]
OC   Mammalia; Eutheria; Rodentia; [illegible]
OC   Mesocricetus.
OX   NCBI_TaxID=10036;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=[illegible]124[illegible]; PubMed=[illegible];
RA   [illegible] A., [illegible] M., [illegible]
RA   [illegible];
RT   "Adr[illegible] (StAR) protein cDNA of hamster adrenals.";
RL   [illegible]
CC   -!- FUNCTION: PLAYS A KEY ROLE IN STEROID HORMONE SYNTHESIS BY
CC       ENHANCING THE METABOLISM OF CHOLESTEROL INTO PREGNENOLONE.
CC       TRANSPORTER THAT BINDS [illegible] AND TRANSFERS [illegible] THE
CC       INTERMEMBRANE SPACE OF THE MITOCHONDRION (PROBABLE).
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL.
CC   -!- SIMILARITY: BELONGS TO THE STAR FAMILY.
CC
CC   This SWISS-PROT entry is copyright. It is produced through [illegible]
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions [illegible]
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. [illegible] commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to [illegible]).
CC
DR   EMBL; [illegible]; AAB[illegible]
DR   InterPro; IPR[illegible];
DR   InterPro; IPR[illegible];
DR   Pfam; [illegible]; START; [illegible]
DR   PRINTS; [illegible]; STARPROTEIN.
KW   Lipid-binding; [illegible] Transport; [illegible]
KW   Mitochondrion; Transit peptide.
FT   TRANSIT       [illegible]       MITOCHONDRION (POTENTIAL).
FT   CHAIN         [illegible]    284       STEROIDOGENIC ACUTE REGULATORY PROTEIN.
SQ   SEQUENCE   [illegible] AA;  [illegible] MW;  [illegible]

  Query Match          [illegible]
  Best Local Similarity   [illegible];  Pred. No. [illegible];
  Matches   [illegible]

Qy          [illegible] MV [illegible]

Db          [illegible] MV [illegible]


RESULT [illegible]
STAR_BOVIN
ID   STAR_BOVIN     STANDARD;      PRT;   [illegible] AA.
AC   [illegible];
DT   [illegible] (Rel. [illegible], Created)
DT   [illegible] (Rel. [illegible], Last sequence update)
DT   [illegible]-MAY-[illegible] (Rel. [illegible], Last annotation update)
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN, [illegible]
GN   STAR.
OS   [illegible]
OC   Eukaryota; Metazoa; [illegible]
OC   Mammalia; Eutheria; [illegible]
OC   [illegible]
OX   NCBI_TaxID=[illegible];
RN   [illegible]
RP   SEQUENCE FROM N.A.
RC   [illegible]
RX   MEDLINE=[illegible]; PubMed=[illegible];
RA   [illegible] W., [illegible] M., [illegible]
RT   "M[illegible] expression [illegible]
RT   [illegible] protein."
RL   [illegible]
RN   [2]
RP   REVISIONS [illegible] AND [illegible]
RA   [illegible] W., [illegible] M., [illegible]
RL   Submitted (MAY-[illegible]) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: PLAYS A KEY ROLE IN STEROID HORMONE SYNTHESIS BY
CC       ENHANCING THE METABOLISM OF CHOLESTEROL INTO PREGNENOLONE.
CC       [illegible] AND TRANSFERS [illegible] THROUGH THE
CC       INTERMEMBRANE SPACE OF THE MITOCHONDRION (PROBABLE).
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL.
CC   -!- TISSUE SPECIFICITY: [illegible] AND ADRENAL [illegible]
CC   -!- SIMILARITY: BELONGS TO THE STAR FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement [illegible]
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; Y17259; CAA76717.1; -.
DR   EMBL; Y17260; CAA76718.1; -.
DR   InterPro; IPR000799; -.
DR   InterPro; IPR002913; -.
DR   Pfam; PF01852; START; 1.
DR   PRINTS; PR00978; STARPROTEIN.
KW   Lipid-binding; Lipid transport; Transport; Steroidogenesis;
KW   Mitochondrion; Transit peptide.
FT   TRANSIT       1      ?       MITOCHONDRION (POTENTIAL).
FT   CHAIN         ?    285       STEROIDOGENIC ACUTE REGULATORY PROTEIN.
FT   CONFLICT     75     75       H -> Y (IN GENOMIC SEQUENCE).
     CONFLICT     99     99       E -> K (IN GENOMIC SEQUENCE).
     SEQUENCE   285 AA;  31845 MW;  EA[illegible] CRC64;

  Query Match             90.0%;  Score [illegible];  DB [illegible];  Length 285;
  Best Local Similarity   92.9%;  Pred. No. [illegible];
  Matches   13;  Conservative    0;  Mismatches    1;  Indels    [illegible];  Gaps    [illegible];

Qy          1 AEHGPICMVLHPLA 14
              |||||||||| |||
Db        218 AEHGPICMVLRPLA 231


RESULT    8
PPA5_HUMAN
ID   PPA5_HUMAN     STANDARD;      PRT;   323 AA.
AC   P13686;
DT   01-JAN-1990 (Rel. 13, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   TARTRATE-RESISTANT ACID PHOSPHATASE TYPE 5 PRECURSOR (EC [illegible]) (TR
DE   AP) (TARTRATE-RESISTANT ACID ATPASE) (TRATPASE).
GN   ACP5.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
     SEQUENCE FROM N.A.
     TISSUE=Placenta;
     MEDLINE=89079710; PubMed=2909539;
     Ketcham C.M., Roberts R.M., Simmen [illegible] M., Nick H.S.;
RT   "Molecular cloning of the type 5, iron-containing, tartrate-resistant
RT   acid phosphatase from human placenta.";
RL   J. Biol. Chem. 264:557-563(1989).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Placenta;
RX   MEDLINE=90249371; PubMed=2338077;
RA   Lord D.K., Cross N.C.P., Bevilacqua M.A., Rider S.H., Gorman P.A.,
RA   Groves A.V., Moss D.W., Sheer D., Cox T.M.;
RT   "Type 5 acid phosphatase. Sequence, expression and chromosomal
RT   localization of a differentiation-associated protein of the human
RT   macrophage.";
RL   Eur. J. Biochem. 189:287-293(1990).
RN   [3]
RP   SEQUENCE OF 22-37.
RX   MEDLINE=90121201; PubMed=2610619;
RA   Stepan J.J., Lau K.H.W., Mohan S., Kraenzlin M., Baylink D.J.;
RT   "Purification and N-terminal sequence of two tartrate-resistant acid
RT   phosphatases type-5 from the hairy cell leukemia spleen.";
RL   Biochem. Biophys. Res. Commun. 165:1027-1034(1989).
```

```
RN   [4]
RP   SEQUENCE OF 22-41.
RX   MEDLINE=90241255; PubMed=2334436;
RA   Stepan J.J., Lau K.H.W., Mohan S., [illegible], Baylink D.J.;
RT   "Purification and N-terminal amino acid sequence of the tartrate-
RT   resistant acid phosphatase from human osteoclastoma: evidence for a
RT   single structure.";
RL   Biochem. Biophys. Res. Commun. [illegible](1990).
RN   [5]
RP   SEQUENCE OF 22-64 AND 181-200.
RX   MEDLINE=89[illegible]; PubMed=2775[illegible];
RA   Hayman A.R., Warburton M.J., Pringle [illegible] A.S., Cole [illegible], Chambers T.J.;
RT   "Purification and characterization of a tartrate-resistant acid
RT   phosphatase from human osteoclastomas.";
RL   Biochem. J. 26[illegible]:601-609(1989).
RN   [6]
RP   SUBUNITS.
RX   MEDLINE=91336978; PubMed=1872798;
RA   Hayman A.R., Dryden A.J., Chambers T.J., Warburton M.J.;
RT   "Tartrate resistant acid phosphatase from human osteoclastomas is
RT   translated as a single polypeptide.";
RL   Biochem. J. 277:631-634(1991).
CC   -!- CATALYTIC ACTIVITY: AN ORTHOPHOSPHORIC MONOESTER + H(2)O = AN
CC       ALCOHOL + ORTHOPHOSPHATE.
CC   -!- COFACTOR: BINDS TWO IRON ATOMS.
CC   -!- SUBUNIT: EXISTS EITHER AS MONOMER OR, AFTER PROTEOLYTIC
CC       PROCESSING, AS A DIMER OF TWO CHAINS LINKED BY DISULFIDE BOND(S).
CC   -!- SUBCELLULAR LOCATION: LYSOSOMAL.
CC   -!- DISEASE: THIS RELATIVELY MINOR INTRACELLULAR ISOZYME OF ACID
CC       PHOSPHATASE CAN BECOME THE DOMINANT ISOZYME IN CERTAIN
CC       PATHOLOGICAL STATES (GAUCHER'S AND HODGKIN'S DISEASES, THE HAIRY
CC       CELL, THE B CELL, AND THE T CELL LEUKEMIAS).
CC   -!- SIMILARITY: TO OTHER TARTRATE-RESISTANT ACID PHOSPHATASES.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement [illegible]
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; J04430; AAA76849.1; -.
DR   EMBL; X14618; CAA32771.1; -.
DR   PIR; A31812; A31812.
DR   PIR; A33781; A33781.
DR   PIR; B33781; B33781.
DR   PIR; A34638; A34638.
DR   PIR; S05251; S05251.
DR   PIR; S05255; S05255.
DR   PIR; S15752; S15752.
DR   MIM; 171640;
KW   Hydrolase; Glycoprotein; Iron; Lysosome; Signal.
FT   SIGNAL        1     21
FT   CHAIN        22    323       TARTRATE-RESISTANT ACID PHOSPHATASE TYPE
FT                                [illegible]
FT   CARBOHYD    116    116       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD    147    147       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CONFLICT     45     46       AR -> [illegible] (IN REF. [illegible]).
FT   CONFLICT     47     47       E -> [illegible] (IN REF. 5).
FT   CONFLICT     [illegible]     56       N -> [illegible] (IN REF. 5).
FT   CONFLICT    177    178       [illegible] -> DVKL (IN REF. 2).
SQ   SEQUENCE   323 AA;  36357 MW;  [illegible] CRC64;

  Query Match             54.4%;  Score [illegible];  DB [illegible];  Length 323;
  Best Local Similarity   68.8%;  Pred. No. [illegible];
  Matches   11;  Conservative    1;  Mismatches    [illegible];  Indels    [illegible];  Gaps    2;

Qy          1 AEHGPI-CMV-LHPL 13
              |||| [illegible]
Db        210 AEHGPIHC VKQLRPL 225
```

```
RESULT   [illegible]
PPA5_MOUSE
ID   PPA5_MOUSE     STANDARD;      PRT;   [illegible] AA.
AC   Q05117;
DT   01-FEB-1994 [illegible]
DT   01-FEB-199[illegible] [illegible] sequence update)
DT   15-JUL-1999 [illegible] annotation update)
DE   TARTRATE-RESISTANT ACID PHOSPHATASE TYPE [illegible]
DE   AP) (TARTRATE-RESISTANT ACID [illegible]
GN   ACP5 OR TRAP [illegible]
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; [illegible]
OC   Mammalia; Eutheria; [illegible]
OX   NCBI_TaxID=[illegible]
RN   [1]
RP   SEQUENCE FROM N.A.
[illegible] MEDLINE=93[illegible] PubMed=[illegible]
[illegible] Cassady A.I., King [illegible]
[illegible] "Isolation and characterization [illegible]
RT   type 5 acid phosphatase.";
RL   Gene 130:201-[illegible](199[illegible]).
RN   [2]
RP   SEQUENCE OF [illegible] FROM N.A.
RC   TISSUE=Spleen;
RX   MEDLINE=940[illegible] PubMed=[illegible];
RA   Reddy S.V., Scarcez T., Windle [illegible]
RA   Chirgwin J.M., [illegible] Roodman [illegible]
RT   "Cloning and characterization [illegible]
RT   tartrate-resistant acid phosphatase gene.";
RL   J. Bone Miner. Res. [illegible]
CC   -!- FUNCTION: MAY PLAY A ROLE IN THE [illegible]
CC       OSTEOCLASTIC TRA[illegible] ACTS ON NUCLEOTIDE [illegible]
CC       HIGHER AFFINITY, COMPARED WITH OTHER [illegible]
CC   -!- CATALYTIC ACTIVITY: AN ORTHOPHOSPHORIC MONOESTER [illegible]
CC       ALCOHOL + PHOSPHATE.
CC   -!- COFACTOR: BINDS TWO IRON ATOMS.
CC   -!- SUBUNIT: EXISTS EITHER AS MONOMER [illegible]
CC       PROCESSING, AS A DIMER OF TWO CHAINS [illegible]
CC   -!- SUBCELLULAR LOCATION: LYSOSOMAL.
CC   -!- TISSUE SPECIFICITY: CHARACTERISTIC [illegible]
CC   -!- SIMILARITY: HIGH IDENTITY [illegible]
CC       BOVINE SPLEEN PURPLE [illegible] PHOSPHATASE.
CC
CC   This SWISS-PROT entry is copyright. [illegible]
[illegible] between the Swiss Institute of Bioinformatics [illegible]
[illegible] the European Bioinformatics Institute. There are [illegible]
[illegible] use by non-profit institutions as long as [illegible]
CC   modified and this statement is not removed. [illegible]
CC   entities requires a license agreement (See http://www.[illegible]
CC   or send an email to license@isb-sib.ch).
CC
DR   EMBL; M99054; AAA[illegible]
DR   EMBL; M8[illegible]212; AAA[illegible]
DR   PIR; JN0787; JN0[illegible]
DR   MGD; MGI:87883; Acp5.
KW   Hydrolase; Glycoprotein; [illegible]
FT   SIGNAL       [illegible]       BY SIMILARITY.
FT   CHAIN        [illegible]       TARTRATE-RESISTANT ACID [illegible]
FT
FT   CARBOHYD     [illegible]       N-LINKED (GLCNAC...) [illegible]
FT   CARBOHYD     [illegible]       N-LINKED (GLCNAC...) [illegible]
SQ   SEQUENCE   327 AA;  [illegible] MW;  [illegible]

  Query Match              [illegible]
  Best Local Similarity    [illegible]
  Matches   [illegible]  Conservative  [illegible]  Mismatches  [illegible]

Qy          1 AEHGFL[illegible]CMV[illegible]
              ||||||  ||
Db        214 AEHGFLRCFVKN[illegible]
```

```
RESULT   [illegible]
PPA5_RA[illegible]
ID   [illegible]     STANDARD;      PRT;   [illegible] AA.
AC   [illegible]
DT   [illegible] Created)
DT   [illegible] sequence update)
DT   [illegible] annotation update)
DE   TARTRATE-RESISTANT ACID PHOSPHATASE TYPE [illegible]
DE   AP) (TARTRATE-RESISTANT ACID [illegible]
GN   ACP5
OS   Rattus [illegible]
OC   Eukaryota; Metazoa; [illegible]
OC   Mammalia; Eutheria; [illegible] Muridae; Murinae; Rattus.
OX   NCBI_TaxID=[illegible]
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=[illegible]
RX   MEDLINE=[illegible] PubMed=[illegible]
RA   Ek-Rylander B., [illegible] Andersson G.;
RT   "[illegible] type [illegible]
RT   tartrate-resistant acid phosphatase [illegible]";
RL   [illegible]
RN   [2]
RP   SEQUENCE OF [illegible]
RX   MEDLINE=[illegible] PubMed=[illegible]
RA   Ek-Rylander B., [illegible] Andersson G.;
RT   "Characterization of a tartrate-resistant acid phosphatase (ATPase)
RT   [illegible] N-terminal amino acid
RT   sequence.";
RL   [illegible]
CC   -!- FUNCTION: MAY PLAY A ROLE IN THE [illegible] BONE RESORPTION. THE
CC       [illegible] ACTS ON NUCLEOTIDE [illegible] WITH
CC       HIGHER AFFINITY, COMPARED WITH OTHER [illegible]
CC   -!- CATALYTIC ACTIVITY: AN ORTHOPHOSPHORIC MONOESTER [illegible] AN
CC       ALCOHOL + PHOSPHATE.
CC   -!- COFACTOR: BINDS TWO IRON ATOMS.
CC   -!- SUBUNIT: EXISTS EITHER AS MONOMER [illegible]
CC       PROCESSING, AS A DIMER OF TWO CHAINS LINKED BY DISULFIDE BOND(S).
CC   -!- SUBCELLULAR LOCATION: LYSOSOMAL.
CC   -!- TISSUE SPECIFICITY: CHARACTERISTIC [illegible]
CC       [illegible] EXPRESSED IN
CC       [illegible]
CC   -!- SIMILARITY: [illegible] TARTRATE-RESISTANT ACID PHOSPHATASE.
CC
CC   This SWISS-PROT entry is copyright. [illegible]
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.[illegible]
CC   or send an email to license@isb-sib.ch).
CC
DR   EMBL; M[illegible]; AAA[illegible]
DR   PIR; A[illegible]; A[illegible]
KW   Hydrolase; Glycoprotein; [illegible]
FT   SIGNAL       [illegible]
FT   CHAIN        [illegible]       TARTRATE-RESISTANT ACID PHOSPHATASE TYPE
FT
FT   CARBOHYD     [illegible]       N-LINKED (GLCNAC...) [illegible]
FT   CARBOHYD     [illegible]       N-LINKED (GLCNAC...) [illegible]
SQ   SEQUENCE   [illegible] AA;  [illegible] MW;  [illegible]

  Query Match              [illegible]
  Best Local Similarity    [illegible]
  Matches   [illegible]  Mismatches  [illegible]  Gaps  [illegible]

Qy          [illegible] AEH[illegible]CMV[illegible]
              [illegible]
Db        [illegible] AEH[illegible]VKN[illegible]
```

```
RESULT  11
TKTP_ALCEU
ID   TKTP_ALCEU     STANDARD;      PRT;   [illegible] AA.
AC   P21726;
DT   01-MAY-1991 (Rel. 18, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   TRANSKETOLASE, PLASMID (EC 2.2.1.1) (TK).
GN   CBBTP.
OS   Alcaligenes eutrophus (Ralstonia eutropha).
OG   Plasmid megaplasmid pHG1.
OC   Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
OC   Ralstonia.
OX   NCBI_TaxID=510;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=H16 / ATCC 17699;
RX   MEDLINE 94042908; PubMed 8226680;
RA   Schaeferjohann J., Yoo J.-G., Kusian B., Bowien B.;
RT   "The cbb operons of the facultative chemoautotroph Alcaligenes
RT   eutrophus encode phosphoglycolate phosphatase.";
     J. Bacteriol. 175:7329-7340(1993).
     [2]
     SEQUENCE OF 1 77 FROM N.A.
RC   STRAIN-H16 / ATCC 17699;
RX   MEDLINE 90152372; PubMed 2559876;
RA   Kossmann J., Klintworth R., Bowien B.;
RT   "Sequence analysis of the chromosomal and plasmid genes encoding
RT   phosphoribulokinase from Alcaligenes eutrophus.";
RL   Gene 85:247-252(1989);
CC   -!- CATALYTIC ACTIVITY: SEDOHEPTULOSE 7 PHOSPHATE + D-GLYCERALDEHYDE
CC       3-PHOSPHATE - D-RIBOSE 5-PHOSPHATE + D XYLULOSE 5-PHOSPHATE.
CC   -!- COFACTOR: THIAMINE PYROPHOSPHATE.
CC   -!- PATHWAY: CALVIN CYCLE.
CC   -!- SIMILARITY: BELONGS TO THE TRANSKETOLASE FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement [illegible]
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; M68905; AAA20194.1; -.
DR   EMBL; M33562; -; NOT_ANNOTATED_CDS.
DR   PIR; PQ0045; PQ0045.
DR   HSSP; P22254; 1AY0.
DR   InterPro; IPR000360; -.
DR   Pfam; PF00456; transketolase; 1.
     PROSITE; PS00801; TRANSKETOLASE_1; 1.
     PROSITE; PS00802; TRANSKETOLASE_2; 1.
KW   Transferase; Thiamine pyrophosphate; Calvin cycle; Plasmid.
SQ   SEQUENCE   670 AA;  71474 MW;  A[illegible]327627A CRC64;

  Query Match             53.8%;  Score [illegible];  DB 1;  Length 670;
  Best Local Similarity   77.8%;  Pred. No. 5.[illegible];
  Matches    7;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        2 EHGPTCMVL 10
            |:||||:||
Db      516 ENGPTCLVL 524


RESULT  12
LIP4_CANRU
ID   LIP4_CANRU     STANDARD;      PRT;   549 AA.
AC   P32948;
DT   01-OCT-1993 (Rel. 27, Created)
DT   01-OCT-1993 (Rel. 27, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   LIPASE 4 PRECURSOR (EC 3.1.1.3).
GN   LIP4.
OS   Candida rugosa (Yeast) (Candida cylindracea)
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX   NCBI_TaxID=5481;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN ATCC 14830;
RX   MEDLINE=93178975; PubMed=8440480;
RA   Lotti M., Grandori R., Fusetti F., Longhi S., Brocca S.,
RA   Tramontano A., Alberghina L.;
RT   "Cloning and analysis of Candida cylindracea lipase sequences."
RL   Gene 124:45-55(1993).
RN   [2]
RP   REVIEW.
RX   MEDLINE=98451816; PubMed 9778794;
RA   Benjamin S., Pandey A.;
RT   "Candida rugosa lipases: molecular [illegible] and versatility in
RT   biotechnology.";
RL   Yeast 14:1069-1087(1998).
CC   -!- CATALYTIC ACTIVITY: TRIACYLGLYCEROL + H(2)O = DIACYLGLYCEROL +
CC       A FATTY ACID ANION.
CC   -!- SIMILARITY: BELONGS TO THE TYPE-B CARBOXYLESTERASE/LIPASE FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement [illegible]
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; X66007; CAA46806.1; -.
DR   PIR; JN0552; JN0552.
DR   HSSP; P32947; 1CLE.
DR   InterPro; IPR002018; -.
DR   Pfam; PF00135; COesterase; 1.
DR   PROSITE; PS00122; CARBOXYLESTERASE_B_1; [illegible]
DR   PROSITE; PS00941; CARBOXYLESTERASE_B_2; [illegible]
KW   Hydrolase; Lipid degradation; Signal; Glycoprotein; Multigene family.
FT   SIGNAL        1     15
FT   CHAIN        [illegible]    [illegible]       LIPASE 4.
FT   ACT_SITE    224    224       BY SIMILARITY.
FT   ACT_SITE    356    356       BY SIMILARITY.
FT   ACT_SITE    464    464       BY SIMILARITY.
FT   DISULFID     [illegible]    112       BY SIMILARITY.
FT   DISULFID    283    292       BY SIMILARITY.
FT   CARBOHYD     66    366       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   549 AA;  58570 MW;  [illegible] CRC64;

  Query Match             52.5%;  Score [illegible];  DB 1;  Length 549;
  Best Local Similarity   54.5%;  Pred. No. [illegible];
  Matches    6;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy        3 HGPTCMVLBEL 13
             :||:|| ::||
Db       71 YGPSCMQMNPL 81


RESULT  13
NEUM_CHICK
ID   NEUM_CHICK     STANDARD;      PRT;   246 AA.
AC   P35001;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   NEUROMODULIN (AXONAL MEMBRANE PROTEIN GAP-43) (PP46) (B-50) (PROTEIN
DE   F1) (CALMODULIN BINDING PROTEIN [illegible]).
GN   GAP43.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=[illegible];
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=901[illegible]; PubMed=21[illegible];
RA   Baizer L., Alkan S., [illegible]
RT   "Chicken growth-associated protein [illegible]
RT   regulated expression of mRNA during embryogenesis."
RL   Brain Res. Mol. Brain Res. [illegible]
CC   -!- FUNCTION: [illegible]
CC       MAJOR COMPONENT OF THE MOTILE "GROWTH CONES" [illegible]
CC       OF ELONGATING AXONS.
CC   -!- SUBCELLULAR LOCATION: [illegible]
CC       SYNAPTIC PLASMA MEMBRANES.
CC   -!- PTM: PHOSPHORYLATION OF THIS PROTEIN BY A [illegible] KINASE [illegible]
CC       SPECIFICALLY [illegible] ATED WITH CERTAIN [illegible]
CC   -!- MISCELLANEOUS: BINDS CALMODULIN WITH A [illegible] IN THE
CC       ABSENCE OF CA[illegible] THAN IN ITS PRESENCE.
CC   -!- SIMILARITY: BELONGS TO THE NEUROMODULIN FAMILY.
CC   -!- SIMILARITY: CONTAINS 1 IQ DOMAIN.
DR   PIR; A6004[illegible]; A[illegible]49.
DR   InterPro; IPR00[illegible]18;
DR   InterPro; IPR001122;
DR   Pfam; PF0061[illegible]
DR   PRINTS; PR00[illegible]; NEUROMODULIN.
DR   PROSITE; PS50096; IQ; 1.
DR   PROSITE; PS00412; NEUROMODULIN_1; 1.
DR   PROSITE; PS00413; NEUROMODULIN_2; 1.
KW   Neurone; Phosphorylation; Membrane; [illegible]
KW   Calmodulin-binding; Lipoprotein; Palmitate.
FT   DOMAIN       [illegible]        [illegible]       [illegible] MEMBRANE [illegible]
FT   DOMAIN       [illegible]        [illegible]       IQ.
FT   LIPID        [illegible]        [illegible]       PALMITATE (BY SIMILARITY).
FT   LIPID        [illegible]        [illegible]       PALMITATE (BY SIMILARITY).
FT   MOD_RES      41                 [illegible]       PHOSPHORYLATION [illegible]
SQ   SEQUENCE   2[illegible] AA;  [illegible] MW;  [illegible]

  Query Match             [illegible];  Score [illegible];  DB [illegible];  Length [illegible];
  Best Local Similarity   [illegible];  Pred. No. [illegible];
  Matches [illegible];  Conservative [illegible];  Mismatches [illegible];  Indels [illegible];  Gaps [illegible];

Qy        [illegible] MVLHEIA [illegible]
              [illegible]
Db    177 [illegible]LLLIPLA [illegible]

RESULT 14
PPA5_PIG
ID   PPA5_PIG       STANDARD;      PRT;   [illegible]4 AA.
AC   P09889;
DT   01-MAR-1989 (Rel. 10, Created)
DT   01-OCT-1993 (Rel. 27, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   TARTRATE-RESISTANT ACID PHOSPHATASE TYPE 5 PRECURSOR [illegible]
DE   AP) (TARTRATE-RESISTANT ACID ATPASE) (TRA[illegible]
GN   ACP5.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; [illegible]
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=98[illegible];
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC   TISSUE=Spleen;
RX   MEDLINE=93216[illegible]; PubMed=[illegible];
RA   Ling P., Roberts R.M.;
RT   "Uteroferrin and intracellular tartrate-resistant acid phosphatases
RT   are the products of the same gene.";
RL   J. Biol. Chem. 268:6896-[illegible]
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=[illegible]; PubMed=[illegible];
RA   Simmen R.C.M., Srinivas V., Roberts R.M.;
```

```
RT   "[illegible]NA sequence [illegible] mRNA
RT   [illegible]"
RL   [illegible]NA [illegible]
RN   [illegible]
RP   SEQUENCE [illegible] N.A.
RX   MEDLINE=[illegible]; PubMed=[illegible]
RA   [illegible] M., [illegible] A., Roberts R.M.;
RT   "M[illegible] and temporal expression [illegible] preparations of the
RT   [illegible]
RT   [illegible]"
RL   [illegible]
RN   [illegible]
RP   SEQUENCE [illegible] N.A.
RX   MEDLINE=[illegible]; PubMed=[illegible]
RA   Yasuda [illegible] H., Mikami H.
RT   "Assignment of the [illegible] acid phosphatase gene (ACP[illegible]) to swine chromosome [illegible]
RT   [illegible]"
RL   Cytogenet. Cell Genet. [illegible]
RN   [illegible]
RP   PARTIAL SEQUENCE.
RX   MEDLINE=[illegible]; PubMed=[illegible]
RA   [illegible] Yates [illegible] Dutta[illegible] [illegible]
RA   A[illegible] Stewart [illegible] Roberts R.M.,
RA   Brew K.
RT   "[illegible] the metalloprotein [illegible] purple acid phosphatase
RT   [illegible] uteroferrin [illegible]"
RL   [illegible]
CC   -!- FUNCTION: [illegible] PROTEIN IS A PHOSPHOPROTEIN PHOSPHATASE [illegible]
CC       [illegible] IN [illegible]
CC       [illegible]
CC   -!- CATALYTIC ACTIVITY: [illegible] ORTHOPHOSPHORIC MONOESTER + H(2)O = AN
CC       [illegible]
CC   -!- [illegible]
CC   -!- SUBCELLULAR LOCATION: SECRETED.
CC   -!- SIMILARITY: [illegible] TARTRATE-RESISTANT ACID PHOSPHATASES.
CC
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC
DR   EMBL; M[illegible]; AAA[illegible]
DR   EMBL; M[illegible]; AAA[illegible]
DR   EMBL; M[illegible]; AAA[illegible]
DR   EMBL; M[illegible]; AAA[illegible]
DR   EMBL; M[illegible]; AAA[illegible]
DR   EMBL; [illegible]; AA[illegible]
DR   PIR; A[illegible]; A[illegible]
DR   PIR; A[illegible]; A[illegible]
KW   Hydrolase; Glycoprotein; [illegible] transport; Signal.
FT   SIGNAL       [illegible]        [illegible]
FT   CHAIN        [illegible]        [illegible]       TARTRATE-RESISTANT ACID PHOSPHATASE TYPE
FT                                                     [illegible]
FT   CARBOHYD     [illegible]        [illegible]       N-LINKED (GLCNAC...) [illegible]
FT   CARBOHYD     [illegible]        [illegible]       N-LINKED (GLCNAC...) [illegible] (POTENTIAL
FT   [illegible]  [illegible]        [illegible]       [illegible]
SQ   SEQUENCE   [illegible] AA;  [illegible] MW;  [illegible]

  Query Match             [illegible]
  Best Local Similarity   [illegible];  Pred. No. [illegible];
  Matches [illegible];  Conservative [illegible];  Mismatches [illegible];  Indels [illegible];  Gaps [illegible];

Qy        [illegible] MV [illegible]
              [illegible]
Db    [illegible] [illegible]

[illegible]
[illegible]
```

```
ID   PRL2_ARATH     STANDARD;      PRT;   479 AA.
AC   Q39190;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   PP1/PP2A PHOSPHATASES PLEIOTROPIC REGULATOR PRL2.
GN   PRL2.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC   Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. COLUMBIA;
RX   MEDLINE=98438452; PubMed=9765207;
RA   Nemeth K., Salchert K., Putnoky P., Bhalerao R., Koncz-Kalman Z.,
RA   Stankovic-Stangeland B., Bako L., Mathur J., Okresz L., Stabel S.,
RA   Geigenberger P., Stitt M., Redei G.P., Schell J., Koncz C.;
RT   "Pleiotropic control of glucose and hormone responses by PRL1, a
RT   nuclear WD protein, in Arabidopsis.";
RL   Genes Dev. 12:3059-3073(1998).
CC   -!- FUNCTION: PLEIOTROPIC REGULATOR OF PP1 AND PP2A PHOSPHATASES.
CC   -!- SIMILARITY: CONTAINS 7 WD REPEATS (TRP-ASP DOMAINS).
CC   -!- SIMILARITY: BELONGS TO THE PRL1/PRL2 FAMILY OF WD REPEAT PROTEINS.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; X82826; CAA58033.1; -.
DR   InterPro; IPR001680; -.
DR   Pfam; PF00400; WD40; 7.
DR   PRINTS; PR00320; GPROTEINBRPT.
DR   PROSITE; PS00678; WD_REPEATS_1; 2.
DR   PROSITE; PS50082; WD_REPEATS_2; 4.
DR   PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW   Repeat; WD repeat.
FT   REPEAT      168    198       WD 1.
FT   REPEAT      210    240       WD 2.
FT   REPEAT      252    282       WD 3.
FT   REPEAT      293    323       WD 4.
FT   REPEAT      335    364       WD 5.
FT   REPEAT      377    406       WD 6.
FT   REPEAT      426    456       WD 7.
SQ   SEQUENCE   479 AA;  53568 MW;  30198A44EA80670 CRC64;

  Query Match             50.0%;  Score 4;  DB 1;  Length 479;
  Best Local Similarity   60.0%;  Pred. No. 5;
  Matches    6;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          3 HGPTCMVLHP 12
              ||  |: |||
Db        255 HGVYCLALHP 264


Search completed: August 22, 2001, 14:30:33
Job time: 407 sec
```

```
                         GenCore version [illegible].5
                Copyright (c) 1993 - 20[illegible] Compugen Ltd.

OM protein - protein search, using sw model

Run on:          August 22, 2001, 14:27:[illegible]   Search time 112.12 Seconds
                                                      (without alignments)
                                                      30.520 Million cell updates/sec

Title:           US-08-659-254-3
Perfect score:   80
Sequence:        1 AEHGPTCMVLHPLA 14

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        425026 seqs, 12230502[illegible] residues

Total number of hits satisfying chosen parameters:       425026

[illegible]mum DB seq length: 0
[illegible]mum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       SPTREMBL_16:*
                 1:  sp_archea:*
                 2:  sp_bacteria:*
                 3:  sp_fungi:*
                 4:  sp_human:*
                 5:  sp_invertebrate:*
                 6:  sp_mammal:*
                 7:  sp_mhc:*
                 8:  sp_organelle:*
                 9:  sp_phage:*
                 10:  sp_plant:*
                 11:  sp_rodent:*
                 12:  sp_unclassified:*
                 13:  sp_vertebrate:*
                 14:  sp_virus:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                      SUMMARIES
                   %
[illegible]lt           Query
 No.   Score  Match Length DB  ID                              Description
-------------------------------------------------------------------------------
   1      72   90.0    285  6  Q9GMD0                          q9gmd0 ovis aries
   2      65   81.2    281 13  Q9DG09                          q9dg09 gallus gal.
   3      61   76.2    289 13  Q9DG08                          q9dg08 xenopus lae
   4      53   66.2    285 13  Q9DG10                          q9dg10 brachydanio
   5      53   66.2    287 13  Q9[illegible]EB3                [illegible] oncorhync[illegible]
   6      53   66.2    287 13  Q9DE06                          q9de06 salvelinus
   7      45   56.2    730  2  Q9RWG0                          q9rwg0 deinococcus
   8    44.5   55.6    269  2  Q9K8U6                          q9k8u6 bacillus ha
   9    43.5   54.4    325  6  O97860                          o97860 oryctolagus
  10    41.5   51.9    325  5  O46177                          o46177 [illegible]
  11      41   51.2    317  2  Q9I200                          q9i200 pseudomonas
  12      41   51.2    442  2  Q9HTM8                          q9htm8 pseudomonas
  13      41   51.2    515  2  Q9Z6U9                          q9z6u9 chlamydia p
  14      41   51.2    515  2  Q9JS84                          q9js84 chlamydia [illegible]
  15      41   51.2    638 10  Q9LQF8                          q9lqf8 arabidopsis
  16      41   51.2    665  5  Q9VPZ5                          q9vpz5 drosophila
  17      41   51.2    962 13  Q9W6N7                          q9w6n7 tinamus gut
  18      41   51.2   1849  4  O14572                          o14572 homo sapien
  19      41   51.2   3012  4  Q9Y3N6                          q9y3n6 homo sapien
  20      40   50.0    206 14  Q9IQY3                          q9iqy3 human immun
  21      40   50.0    396 10  Q9MAH8                          q9mah8 arabidopsis
  22      40   50.0    453 10  O48772                          o48772 arabidopsis
  23      40   50.0    476 10  Q9LJR9                          q9ljr9 arabidopsis
  24      40   50.0    497  2  Q9EW37                          q9ew37 streptomyce
  25      40   50.0    684  2  Q9RVA3                          q9rva3 deinococcus
  26      39   48.8    228  5  Q9XWI1                          q9xwi1 caenorhabdi
  27      39   48.8    372  2  Q9KT15                          q9kt15 vibrio chol
  28      39   48.8    413 10  Q9[illegible]                   [illegible] oryza sativ
  29      39   48.8    459  2  O25796                          o25796 helicobacte
  30      39   48.8    459  2  Q9ZK82                          q9zk82 helicobacte
  31      39   48.8    483 13  Q9IAB0                          q9iab0 agkistrodon
  32      39   48.8    512  2  O84814                          o84814 chlamydia t
  33      39   48.8    512  2  Q9PLB4                          q9plb4 chlamydia m
  34      39   48.8    614  5  Q9G[illegible]                  [illegible] drosophila
  35      39   48.8    754  5  Q9W[illegible]                  [illegible] drosophila
  36      39   48.8    816  5  Q9V[illegible]                  [illegible] drosophila
  37      39   48.8   1309 10  Q9MAJ8                          q9maj8 arabidopsis
  38      39   48.8   1872  5  Q17[illegible]                  [illegible]
  39    38.5   48.1   1484  5  Q9N8[illegible]                 [illegible] trypanosoma
  40      38   47.5    208 11  O54[illegible]                  [illegible] mus musculu
  41      38   47.5    240  5  Q9QX[illegible]                 [illegible] leishmania
  42      38   47.5    243  6  O19[illegible]                  [illegible] bos taurus
  43      38   47.5    269  4  O95[illegible]                  [illegible] homo sapien
  44      38   47.5    281 10  [illegible]                     [illegible] arabidopsis
  45      38   47.5    323  5  Q9VA[illegible]                 [illegible] drosophila

                                     ALIGNMENTS

RESULT      1
Q9GMD0
ID   Q9GMD0      PRELIMINARY;      PRT;   285 AA.
AC   Q9GMD0;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN.
GN   STAR.
OS   Ovis aries (Sheep).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Caprinae; Ovis.
OX   NCBI_TaxID 9940;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   West L.A., Horvat R.D., Roess D.A., Barisas B.G., Juengel J.[illegible]
RA   Niswender G.D.;
RT   "Steroidogenic Acute Regulatory Protein and Peripheral-type
RT   Benzodiazepine Receptor Associate at the Mitochondrial Membrane.";
RL   Submitted (JUL 2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF290202; AAG02464.1; -.
SQ   SEQUENCE   285 AA;  31895 MW;  A8[illegible] CRC64;

  Query Match             90.0%;  Score 72;  DB 6;  Length 285;
  Best Local Similarity   92.9%;  Pred. No. [illegible];
  Matches   13;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 AEHGPTCMVLHPLA 14
              ||||||||||| ||
Db        218 AEHGPTCMVLRPLA 231


RESULT      2
Q9DG09
ID   Q9DG09      PRELIMINARY;      PRT;   281 AA.
AC   Q9DG09;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN.
```

```
GN   STAR
OS   Gallus gallus [illegible]
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; [illegible]
OC   Archosauria; Aves; Neognathae; [illegible]
OC   Gallus.
OX   NCBI_TaxID [illegible]
RN   [1]
RP   SEQUENCE FROM N.A.
RX   PubMed 110641[illegible]
RA   Bauer M.P., Bridgham [illegible], Langenau [illegible]
RT   "Conservation of steroidogenic acute regulatory [illegible]
RT   structure and expression in vertebrates.";
RL   Mol. Cell. Endocrinol. 168:11[illegible](2000).
DR   EMBL; AF2204[illegible]; AAG[illegible]
SQ   SEQUENCE   [illegible] AA;  [illegible] MW;  [illegible]

  Query Match             [illegible]
  Best Local Similarity   [illegible]
  Matches   12;  Conservative  [illegible]  Mismatches  [illegible]

Qy          1 AEHGP[illegible]MV[illegible]PLA [illegible]
                [illegible]
Db        218 AENGP[illegible]MV[illegible]PLA [illegible]


RESULT [illegible]
Q9DG08
ID   Q9DG08      PRELIMINARY;      PRT;   [illegible] AA.
AC   Q9DG08;
DT   01-MAR-2001 (TrEMBLrel. [illegible], Created)
DT   01-MAR-2001 (TrEMBLrel. [illegible], Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. [illegible], Last annotation update)
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN (FRAGMENT).
GN   STAR
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; [illegible]
OC   Amphibia; Batrachia; Anura; Mesobatrachia; [illegible]
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=OVARY;
RX   PubMed 110641[illegible]
RA   Bauer M.P., Bridgham J.T., Langenau D.M., [illegible]
RT   "Conservation of steroidogenic acute regulatory [illegible]
RT   structure and expression in vertebrates.";
RL   Mol. Cell. Endocrinol. 168:119-1[illegible](2[illegible]
DR   EMBL; AF2204[illegible]; AAG2[illegible]
FT   NON_TER  [illegible]
FT   NON_TER  [illegible]
SQ   SEQUENCE   [illegible] AA;  [illegible] MW;  [illegible]

  Query Match             [illegible]
  Best Local Similarity   [illegible]
  Matches   11;  Conservative  [illegible]  Mismatches  [illegible]

Qy          2 EHGPTCMV[illegible]HLA [illegible]
                | |||||||
Db        223 ENGPTCMV[illegible]PLA [illegible]


RESULT 4
Q9DG10
ID   Q9DG10      PRELIMINARY;      PRT;   [illegible] AA.
AC   Q9DG10;
DT   01-MAR-2001 (TrEMBLrel. [illegible], Created)
DT   01-MAR-2001 (TrEMBLrel. [illegible], Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. [illegible], Last annotation update)
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN.
GN   STAR
OS   Brachydanio rerio [illegible]
```

```
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; [illegible]
OC   Cypriniformes; [illegible]
OX   NCBI_TaxID [illegible]
RN   [illegible]
RP   SEQUENCE FROM N.A.
RX   PubMed [illegible]
RA   [illegible] Langenau D.M., [illegible] A., [illegible] E.W.;
RT   "[illegible] (StAR) protein
RT   structure and expression in vertebrates.";
RL   [illegible]
DR   EMBL; AF[illegible]
SQ   SEQUENCE   [illegible] MW;  [illegible]

  Query Match             [illegible]  Length [illegible]
  Best Local Similarity   [illegible]  Pred. No. [illegible]
  Matches   [illegible]  Conservative  [illegible]  Mismatches  [illegible]  Indels  [illegible]  Gaps  [illegible]

Qy          [illegible] A[illegible]MV[illegible]A [illegible]

Db        [illegible] AEN[illegible]VM[illegible]A [illegible]


RESULT [illegible]
[illegible]
ID   [illegible]      PRELIMINARY;      PRT;   [illegible]
AC   [illegible]
DT   [illegible] MAR [illegible] Created)
DT   [illegible] MAR [illegible] Last sequence update)
DT   [illegible] MAR [illegible] Last annotation update)
DE   [illegible] ACUTE [illegible] PROTEIN.
GN   STAR
OS   [illegible]
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; [illegible]; Teleostei;
OC   [illegible]
OX   NCBI_TaxID [illegible]
RN   [illegible]
RP   SEQUENCE FROM N.A.
RC   [illegible] AND [illegible]
RA   [illegible] Kusakabe M., [illegible]
RT   "Molecular [illegible] (StAR) [illegible]
RT   [illegible]";
RL   Submitted [illegible] to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB[illegible]
SQ   SEQUENCE   [illegible] AA;  [illegible] MW;  [illegible]

  Query Match             [illegible]
  Best Local Similarity   [illegible]  Pred. No. [illegible]
  Matches   [illegible]  Conservative  [illegible]  Mismatches  [illegible]  Indels  [illegible]  Gaps  [illegible]

Qy          [illegible] AEH[illegible]MV[illegible]

Db        [illegible] AEN[illegible]VM[illegible]


RESULT [illegible]
[illegible]
ID   [illegible]      PRELIMINARY;      PRT;   [illegible] AA.
AC   [illegible]
DT   [illegible] MAR [illegible] (TrEMBLrel. [illegible] Created)
DT   [illegible] MAR [illegible] (TrEMBLrel. [illegible] Last sequence update)
DT   [illegible] MAR [illegible] (TrEMBLrel. [illegible] Last annotation update)
DE   [illegible] N [illegible] A [illegible] PROTEIN.
GN   STAR
OS   [illegible]
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; [illegible]; Euteleostei;
OC   [illegible]
OX   NCBI_TaxID [illegible]
RN   [illegible]
```

```
RP   SEQUENCE FROM N.A.
RC   TISSUE BRAIN, TESTIS, HEAD, KIDNEY [illegible] AND LIVER;
RA   Goetz F.W.;
RT   "Characterization of the brook trout [illegible] and [illegible] homologs.";
RL   Submitted (FEB-2000) to the EMBL/[illegible] databases.
DR   EMBL; AF232215; AAG39689.1; -.
SQ   SEQUENCE   287 AA;  31884 MW;  29F[illegible]4EFAF 5F[illegible]E9B CRC64;

  Query Match             66.2%;  Score 53;  DB 13;  Length 287;
  Best Local Similarity   64.3%;  Pred. No. [illegible];
  Matches    9;  Conservative    [illegible];  Mismatches    2;  Indels    [illegible];  Gaps    [illegible];

Qy        1 AEHGPTCMVLHPLA 14
            ||:||||:|: | |
Db      217 AENCPTCIVMRPSA 230


RESULT     7
Q9RWG0
ID   Q9RWG0       PRELIMINARY;       PRT;   [illegible] AA.
     Q9RWG0;
     01-MAY-2000 (TrEMBLrel. 13, Created)
     01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   HYPOTHETICAL 78.1 KDA PROTEIN.
GN   DR0708.
OS   Deinococcus radiodurans.
OC   Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.
OX   NCBI_TaxID=1299;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=R1;
RX   MEDLINE=20036896; PubMed=10567266;
RA   White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA   Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA   Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA   Vamathevan J.J., Lam P., McDonald L., [illegible], [illegible],
RA   Makarova K.S., Aravind L., Daly M.J., [illegible] K.W., [illegible],
RA   Ketchum K.A., Nelson K.E., [illegible], [illegible] H., [illegible],
RA   Fraser C.M.;
RT   "Genome sequence of the radioresistant bacterium Deinococcus
RT   radiodurans R1.";
RL   Science 286:1571-1577(1999).
DR   EMBL; AE001927; AAF10291.1; -.
DR   TIGR; DR0708; -.
DR   InterPro; IPR000508; -.
DR   PROSITE; PS00761; SPASE_I_3; UNKNOWN_1.
KW   Hypothetical protein.
SQ   SEQUENCE   730 AA;  78073 MW;  9AE6[illegible]30428AA5 CRC64;

 Query Match             56.2%;  Score 4[illegible];  DB 2;  Length 730;
  Best Local Similarity   88.9%;  Pred. No. [illegible];
  Matches    8;  Conservative    0;  Mismatches    1;  Indels    [illegible];  Gaps    [illegible];

Qy        5 PTCMVLHPL 13
            ||| |||||
Db      536 PTCGVLHPL 544


RESULT     8
Q9K8U6
ID   Q9K8U6       PRELIMINARY;       PRT;   [illegible] AA.
AC   Q9K8U6;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   BH2906 PROTEIN.
GN   BH2906.
OS   Bacillus halodurans.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group;
OC   Bacillus/Staphylococcus group; Bacillus.
```

```
OX   NCBI_TaxID=86665;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C-125 / JCM 9153;
RA   Takami H., [illegible] K., Takaki Y.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AP001517; BAB06625.1; -.
DR   InterPro; IPR000150; -.
DR   InterPro; IPR00[illegible]454; -.
DR   Pfam; PF00702; Hydrolase; 1.
DR   PROSITE; PS01228; [illegible]; 1.
SQ   SEQUENCE   26[illegible] AA;  30291 MW;  17A[illegible]4CC161E[illegible]74 CRC64;

  Query Match             55.6%;  Score 43.5;  DB 2;  Length 26[illegible];
  Best Local Similarity   40.0%;  Pred. No. [illegible];
  Matches    8;  Conservative    3;  Mismatches    0;  Indels    9;  Gaps    1;

Qy        2 EHG---------PTCMVLHP 12
            | |          |:::|.
Db      136 EHGNLLKKELR[illegible]PICVLIHP 155


RESULT     9
O97860
ID   O97860       PRELIMINARY;       PRT;   325 AA.
AC   O97860;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   TARTRATE-RESISTANT ACID PHOSHATASE.
OS   Oryctolagus cuniculus (Rabbit)
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX   NCBI_TaxID=9986;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=JAPAN WHITE, TISSUE=LONG BONES;
RX   MEDLINE=981[illegible]; PubMed=[illegible];
RA   Kubota M., [illegible], Nara H., Kato M., [illegible] M., Nojima H.,
RA   Kawashima H.;
RT   "Large scale isolation of osteoclast specific genes by an improved
RT   method involving the preparation of a subtracted cDNA library.";
RL   Genes Cells 3:459-475(1998).
CC   -!- CATALYTIC ACTIVITY: AN ORTHOPHOSPHORIC MONOESTER + H(2)O = AN
CC       ALCOHOL + ORTHOPHOSPHATE (BY SIMILARITY).
DR   EMBL; AB009340; BAA75919.1; -.
DR   InterPro; IPR000934; -.
DR   InterPro; IPR00[illegible]266; -.
DR   Pfam; PF02227; FA_phosphatase; 1.
KW   Hydrolase; Iron; Zinc.
SQ   SEQUENCE   325 AA;  36624 MW;  [illegible] CRC64;

  Query Match             54.4%;  Score [illegible];  DB 6;  Length 325;
  Best Local Similarity   68.8%;  Pred. No. [illegible];
  Matches   11;  Conservative    1;  Mismatches    1;  Indels    3;  Gaps    2;

Qy        1 AEHGPT-CMV--LHPL 13
            |||||| |:|   | |
Db      212 AEHGPTHCVVKKLQPL 227


RESULT    10
O46177
ID   O46177       PRELIMINARY;       PRT;   3[illegible] AA.
AC   O46177;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   PRE-PROCATHEPSIN L PRECURSOR.
OS   Paragonimus westermani.
OC   Eukaryota; Metazoa; Platyhelminthes; [illegible]; Neodermata;
OC   Trematoda; [illegible]
OC   Paragonimus.
OX   NCBI_TaxID=[illegible]
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Yun D.H., Chung J.Y., Yang H.J., [illegible]
RL   Submitted (SEP-[illegible]) to the EMBL/GenBank/DDBJ [illegible]
DR   EMBL; [illegible]053[illegible]; AA[illegible]
DR   HSSP; P[illegible]
DR   InterPro; IPR[illegible]
DR   InterPro; IPR000668.
DR   Pfam; PF00112; Peptidase_C1; [illegible]
DR   PRINTS; PR00705; PAPAIN.
DR   PROSITE; PS00[illegible]; THIOL_PROTEASE_ASN; [illegible]
DR   PROSITE; PS00[illegible]; THIOL_PROTEASE_CYS; [illegible]
DR   PROSITE; PS00[illegible]; THIOL_PROTEASE_HIS; [illegible]
KW   Hydrolase; Signal; Thiol protease.
FT   SIGNAL        1     [illegible]       POTENTIAL.
FT   CHAIN        [illegible]   [illegible]       [illegible]
SQ   SEQUENCE   [illegible] AA;  [illegible] MW;  [illegible] CRC64;

  Query Match              [illegible]
  Best Local Similarity    [illegible];  Pred. No. [illegible]
  Matches  [illegible];  Conservative  [illegible];  Mismatches  [illegible];  Indels  [illegible];  Gaps  [illegible]

Qy          1 AFHGP[illegible]M[illegible]VIH[illegible]
                [illegible]
Db        236 AFHGPMS[illegible]ENA[illegible]PL[illegible]YEY[illegible]LH[illegible]


RESULT 11
Q9I200
ID   Q9I200     PRELIMINARY;   PRT;   [illegible] AA.
AC   Q9I200;
DT   01-MAR-2001 (TrEMBLrel. [illegible], Created)
DT   01-MAR-2001 (TrEMBLrel. [illegible], Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. [illegible], Last annotation update)
DE   PROBABLE TRANS[illegible]NAL REGULATOR.
GN   PA2115.
OS   Pseudomonas aeruginosa.
OC   Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae; [illegible]
OC   Pseudomonas.
OX   NCBI_TaxID=287;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PAO1;
RX   MEDLINE=2043[illegible]; PubMed=[illegible]84[illegible]
RA   Stover C.K., Pham X.[illegible], Erwin A.L., Mi[illegible], Wa[illegible]
RA   Hickey M.J., Brinkman [illegible], Hufnagle W.[illegible]
RA   Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RA   Brody L.L., Coulter S.N., Folger K.R., Kas A., [illegible]
RA   Smith K.A., Spencer D.[illegible], Wong G.K.[illegible], Wu Z.[illegible]
RA   Reizer J., Saier M.H., Hancock R.E.W., [illegible]
RT   "Complete genome sequence of Pseudomonas aeruginosa PAO1, an
RT   opportunistic pathogen."
RL   Nature 406:959-964(2000).
DR   EMBL; AE004[illegible]; AAG0[illegible]
DR   InterPro; IPR[illegible]
DR   Pfam; PF00126; HTH_1; [illegible]
DR   PRINTS; PR00[illegible]; HTH[illegible]
SQ   SEQUENCE   [illegible] AA;  [illegible] MW;  [illegible] CRC64;

  Query Match              [illegible]
  Best Local Similarity    [illegible];  Pred. No. [illegible]
  Matches    8;  Conservative  [illegible];  Mismatches  [illegible];  Indels  [illegible];  Gaps  [illegible]

Qy          1 AFHGPTCMVLHI[illegible] 13
                [illegible]
Db        [illegible] ANV[illegible]PECMLLAL[illegible] 30
```

```
RESULT 12
Q9[illegible]
ID   [illegible]     PRELIMINARY;   PRT;   [illegible] AA.
AC   [illegible]
DT   01-MAR-[illegible] (TrEMBLrel. [illegible], Created)
DT   01-MAR-[illegible] (TrEMBLrel. [illegible], Last sequence update)
DT   01-MAR-[illegible] (TrEMBLrel. [illegible], Last annotation update)
DE   PROBABLE [illegible]
GN   PA[illegible]
OS   Pseudomonas aeruginosa.
OC   Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae; [illegible]
OC   Pseudomonas.
OX   NCBI_TaxID=287;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PAO1;
RX   MEDLINE=[illegible]; PubMed=[illegible]
RA   Stover C.K., Pham X.[illegible], Erwin A.L., [illegible], Warrener P.,
RA   Hickey M.J., Brinkman F.S.L., Hufnagle W.[illegible], Kowalik D.J., Lagrou M.,
RA   Garber R.L., [illegible] Westbrock-Wadman S., Yuan Y.,
RA   Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA   Smith K.A., Spencer D.H., Wong G.K., [illegible]
RA   Reizer J., Saier M.H., Hancock R.E.W., [illegible] Olson M.V.,
RT   "Complete genome sequence of Pseudomonas aeruginosa PAO1, an
RT   opportunistic pathogen."
RL   Nature 406:959-964(2000).
DR   EMBL; AE[illegible]
DR   InterPro; [illegible]
DR   Pfam; [illegible]
SQ   SEQUENCE   441 AA;  [illegible] MW;  [illegible] CRC64;

  Query Match              [illegible]
  Best Local Similarity    [illegible];  Pred. No. [illegible]
  Matches  [illegible];  Conservative  [illegible];  Mismatches  [illegible];  Indels  [illegible];  Gaps  [illegible]

Qy          [illegible] AFHGP[illegible]M[illegible]

Db        [illegible] AF[illegible]

RESULT 13
Q9Z[illegible]
ID   [illegible]     PRELIMINARY;   PRT;   [illegible] AA.
AC   [illegible]
DT   01-MAY-[illegible] (TrEMBLrel. [illegible], Created)
DT   01-MAY-[illegible] (TrEMBLrel. [illegible], Last sequence update)
DT   [illegible] (TrEMBLrel. [illegible], Last annotation update)
DE   AXIAL FILAMENT PROTEIN ([illegible]N [illegible]LEASE [illegible])
GN   [illegible]
OS   Chlamydophila pneumoniae (Chlamydia pneumoniae).
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX   NCBI_TaxID=[illegible]
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CW[illegible]
RX   MEDLINE=[illegible]; PubMed=[illegible]
RA   Kalman S., Mitchell W., Marathe R., [illegible], Hyman R.W.,
RA   [illegible], Stephens R.S.;
RT   "Comparative genomes of Chlamydia pneumoniae and C. trachomatis."
RL   Nat. Genet. [illegible](1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AR39;
RX   MEDLINE=[illegible]; PubMed=[illegible]
RA   Read T.D., Brunham R.C., [illegible], Heidelberg J.F.,
RA   White O., Hickey E.K., [illegible], Utterback T.,
RA   [illegible], Weidman J., Khouri H., Craven B.,
RA   Bowman C., [illegible], Gwinn M., Nelson W., DeBoy R., Kolonay J.,
RA   McClarty G., Salzberg S.L., [illegible], Fraser C.M.;
RT   "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
RT   pneumoniae AR39.";
RL   Nucleic Acids Res. [illegible](2000).
```

```
DR   EMBL; AE001675; AAD19095.1; -.
DR   EMBL; AE002249; AAF38687.1; -.
DR   TIGR; CP0901; -.
DR   InterPro; IPR003029; -.
DR   Pfam; PF00575; S1; 1.
SQ   SEQUENCE   515 AA;  59550 MW;  A7FE[illegible]A86538 CRC64;

  Query Match             51.2%;  Score 41;  [illegible]  Length 515;
  Best Local Similarity   54.5%;  Pred. No. [illegible]
  Matches    6;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    [illegible]

Qy          2 EHGPICMVLHP 12
              ||   |:|:||
Db        453 EHSHLCLVVHP 463


RESULT 14
Q9JS84
ID   Q9JS84      PRELIMINARY;      PRT;   [illegible] AA.
AC   Q9JS84;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   AXIAL FILAMENT PROTEIN.
GN   CAFE.
OS   Chlamydia pneumoniae (Chlamydophila pneumoniae).
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX   NCBI_TaxID=83558;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=J138;
RX   MEDLINE=20330349; PubMed=10871362;
RA   Shirai M., Hirakawa H., Kimoto M., [illegible] M., Kishi F., [illegible] K.,
RA   Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT   "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RT   from Japan and CWL029 from USA.";
RL   Nucleic Acids Res. 28:2311-2314(2000).
DR   EMBL; AP002548; BAA99167.1; -.
DR   InterPro; IPR003029; -.
DR   Pfam; PF00575; S1; 1.
SQ   SEQUENCE   515 AA;  59609 MW;  929EB2A[illegible]3AF718 CRC64;

  Query Match             51.2%;  Score 41;  [illegible]  Length 515;
  Best Local Similarity   54.5%;  Pred. No. [illegible]
  Matches    6;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    [illegible]

Qy          2 EHGPTCMVLHP 12
              ||   |:|:||
Db        453 EHSHLCLVVHP 463

RESULT 15
Q9LQF8
ID   Q9LQF8      PRELIMINARY;      PRT;   6[illegible] AA.
AC   Q9LQF8;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   F1504.27.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC   Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Ecker J.R.;
RL   Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A
RA   Ecker J.R.;
RL   Submitted (OCT-1999) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Cheuk R., Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,
RA   Khan S., Kim C., Altafi H., Bei B., [illegible] Chiou [illegible] Choi E.,
RA   Conn L., Conway A., Gonzalez A., Hansen N., Howing [illegible] Koo T., Lam B.,
RA   Lee J., Lenz C., Li J., Liu A., [illegible] S., Mulkarsky N.,
RA   Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,
RA   Thaveri A., Toriumi M., Vaysberg M., [illegible] Davis R., Federspiel N.,
RA   Theologis A., Ecker J.;
RL   Submitted (JUN-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC007887; AAF79353.1; -.
DR   InterPro; IPR000345; -.
DR   PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
SQ   SEQUENCE   638 AA;  74113 MW;  5AAC[illegible]AAF17FF CRC64;

  Query Match             51.2%;  Score 41;  [illegible]
  Best Local Similarity   54.5%;  Pred. No. 41;
  Matches    6;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 EHGPICMVLHP 12
              | ||  |::  |
Db         30 EDGPRCLMSHP 40


Search completed: August 22, 2001, 14:2[illegible]
Job time: 196 sec
```

```
                         GenCore version 4.5
              Copyright (c) 1993 - 2001 Compugen Ltd.

OM protein - protein search, using sw model

Run on:         August 22, 2001, 14:25:24 ; Search time 61.49 Seconds
                                            (without alignments)
                                            14.174 Million cell updates/sec

Title:          US-08-659-254-4
Perfect score:  66
Sequence:       1 ALGILNNQEGWK 12

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters:      219241

  nimum DB seq length: 0
   imum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       PIR_68:*
               1:  pir1:*
               2:  pir2:*
               3:  pir3:*
               4:  pir4:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                                SUMMARIES
                      %
Result              Query
  No.      Score    Match Length DB  ID                    Description
--------------------------------------------------------------------------
     1        66    100.0    284  2  A55455                steroidogenic acut
     2        66    100.0    362  2  JC5386                steroidogenic acut
     3        61     92.4    285  2  I38248                steroidogenic acut
     4        55     83.3    285  2  JC4315                steroidogenic acut
     5        48     72.7    783  2  JC5467                cellulase (EC 3.2.
     6        48     72.7    824  2  JC7532                cellulase (EC 3.2.
     7        47     71.2    157  2  S62571                probable ubiquitin
     8        43     65.2    800  2  A29003                cellulase (EC 3.2.
     9        43     65.2    822  2  JT0611                cellulase (EC 3.2.
    10        41     62.1    885  2  S42841                T16G12.1 protein
    11        40     60.6   1463  2  T30290                AAS surface protei
    12        39     59.1    132  2  F82603                hypothetical prote
    13        39     59.1    208  2  F81941                hypothetical prote
    14        39     59.1    260  2  T12393                NADH dehydrogenase
    15        39     59.1    346  2  T10528                probable gibberell
    16        39     59.1    479  2  E86785                glutamate synthase
    17        39     59.1    505  2  C81165                conserved hypothet
    18        39     59.1   1445  2  T15202                hypothetical prote
    19        39     59.1   1711  2  I31317                1,4-beta glucanase
    20        39     59.1   1742  2  T17110                cellulase (EC 3.2.
    21        38     57.6    148  2  F86151                hypothetical prote
    22        38     57.6    171  2  E82293                lipoprotein, signal
    23        38     57.6    181  2  E83436                probable oxidase P
    24        38     57.6    218  2  S56516                hypothetical 24.6K
    25        38     57.6    311  2  S38639                replication protei
    26        38     57.6    372  2  C70540                probable nifU prot
    27        38     57.6    481  2  H69543                6-phospho-beta-glu
    28        38     57.6    557  2  C85569                hypothetical prote
    29        38     57.6    620  2  S64304                hypothetical prote
    30        38     57.6    825  2  JS0174                cellulase (EC 3.2.
    31        38     57.6   1289  2  B72454                conserved hypothet
    32        38     57.6   1943  2  B64796                toxin-like outer m
    33      37.5     56.8    605  2  C86588                hypothetical prote
    34        37     56.1     99  2  I43800                probable transposa
    35        37     56.1    126  2  I50603                ubiquitin conjugat
    36        37     56.1    359  2  JQ1049                glucose glutamine
    37        37     56.1    393  2  JC5634                RNB6 protein - rat
    38        37     56.1    859  1  S06418                3',5'-cyclic-GMP p
    39        37     56.1    859  2  B34611                3',5'-cyclic-GMP p
    40        37     56.1    859  2  S13030                3',5'-cyclic-GMP p
    41        37     56.1    947  2  I49635                mouse dhm1 protein
    42        37     56.1   1511  2  A53561                plectin triple dr
    43        37     56.1   1529  2  S69688                hypothetical prote
    44        37     56.1   3535  2  E83611                probable hemagglut
    45        37     56.1   5627  2  C83339                hypothetical prote

                                ALIGNMENTS

RESULT    1
A55455
steroidogenic acute regulatory protein precursor, mitochondrial - mouse
C;Species: Mus musculus (house mouse)
C;Date: 10-Feb-1995 #sequence_revision 10-Feb-1995 #text_change 08-Oct-1999
C;Accession: A55455
R;Clark, B.J.; Wells, J.; King, S.R.; Stocco, D.M.
J. Biol. Chem. 269, 28314-28322, 1994
A;Title: The purification, cloning, and expression of a novel luteinizing hormone-ind
ry protein (StAR).
A;Reference number: A55455; MUID:95050681
A;Accession: A55455
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-284 <CLA>
A;Cross-references: GB:L36062; GB:L36062; GB:S73839; NID:g1236242; PIDN:AAB94783.1; P
C;Keywords: mitochondrion; phosphoprotein
F;1-47/Domain: transit peptide (mitochondrion) #status predicted <TRP>

  Query Match             100.0%;  Score 66;  DB 2;  Length 284;
  Best Local Similarity   100.0%;  Pred. No. 3.9e-26;
  Matches   12;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ALGILNNQEGWK 12
                ||||||||||||
Db         85 ALGILNNQEGWK 96

RESULT    2
JC5386
steroidogenic acute regulatory protein - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 04-Jun-1997 #sequence_revision 18-Jul-1997 #text_change 07-May-1999
C;Accession: JC5386
R;Lee, H.; Ahn, K.S.; Kwon, H.; Soh, J.
Biochem. Biophys. Res. Commun. 230, 528-532, 1997
A;Title: Nucleotide sequence of rat steroidogenic acute regulatory protein complement
A;Reference number: JC5386; MUID:97167737
A;Accession: JC5386
A;Molecule type: mRNA
A;Residues: 1-362 <LEE>
A;Cross-references: GB:U76419
C;Comment: This protein is a key regulator for the steroidogenesis in acute response t

  Query Match             100.0%;  Score 66;  DB 2;  Length 362;
  Best Local Similarity   100.0%;  Pred. No. 5.6e-26;
  Matches   12;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ALGILNNQEGWK 12
                ||||||||||||
Db        163 ALGILNNQEGWK 1[illegible]


RESULT    [illegible]
I38248
steroidogenic acute regulatory protein - human
C;Species: Homo sapiens (man)
C;Date: 29 May 1998 #sequence_revision [illegible]
C;Accession: I38248; I38800
R;Sugawara, T.; [illegible] A.; Martin, [illegible]
Biochemistry 34, 12[illegible]
A;Title: Structure of the human steroidogenic acute regulatory protein [illegible]
A;Reference number: I38248; MUID:[illegible]
A;Accession: I38248
A;Status: preliminary; translated from GB/EMBL[illegible]
A;Molecule type: DNA
A;Residues: 1-285 [illegible]
[illegible]ss references: EMBL:[illegible]; NID:[illegible]; PIDN:AA[illegible]
[illegible]gawara, T.; Holt, J.A.; [illegible]
Proc. Natl. Acad. Sci. U.S.A. [illegible]
A;Title: Human steroidogenic acute regulatory protein: [illegible]
A;Reference number: I38800; MUID:[illegible]
A;Accession: I38800
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-285 [illegible]
A;Cross-references: EMBL:[illegible]; NID:[illegible]; PIDN:AA[illegible]
C;Genetics:
A;Gene: STAR
A;Cross-references: GDB:[illegible]; MIM:[illegible]
A;Map position: 8p11[illegible]
A;Introns: [illegible]


  Query Match             [illegible]
  Best Local Similarity   [illegible]; Pred. No. [illegible]
  Matches   [illegible]; Conservative [illegible]; Mismatches [illegible]

Qy          1 ALGILNNQEGWK 12
                [illegible]
Db         86 ALGILSNQEGWK 97


RESULT    4
[illegible]315
[illegible]oidogenic acute regulatory protein [illegible]
[illegible]pecies: Bos primigenius taurus (cattle)
C;Date: 29 Nov 1997 #sequence_revision [illegible]
C;Accession: I[illegible]
R;Hartung, [illegible]; Rust, W.; [illegible] M. [illegible]
Biochem. Biophys. Res. Commun. 215, [illegible]
A;Title: Molecular cloning and in vivo expression [illegible]
A;Reference number: [illegible]; MUID:[illegible]
A;Accession: I[illegible]
A;Molecule type: mRNA
A;Residues: 1-285 <HAR>
C;Comment: This protein [illegible]
C;Genetics:
A;Gene: STAR
F;226-264/Region: [illegible]


  Query Match             [illegible]
  Best Local Similarity   [illegible]; Pred. No. [illegible]
  Matches   [illegible]; Conservative [illegible]; Mismatches [illegible]

Qy          1 ALGILNNQEGWK 12
                [illegible]
Db         86 ALGIIKDQEGWK 97


RESULT    5
[illegible]
```

```
[illegible]
N;Alternate names: [illegible] 1,4-beta-[illegible]
C;Species: [illegible]
C;Date: [illegible] #sequence_revision [illegible] #text_change [illegible]
C;Accession: [illegible]
R;Miyazaki, M.; [illegible]
Biosci. Biotechnol. Biochem. [illegible]
A;Title: A [illegible] 1,4-beta-[illegible]
A;Reference number: [illegible]; MUID:[illegible]
A;Accession: [illegible]
A;Molecule type: DNA
A;Residues: [illegible]
A;Cross-references: [illegible]
A;Experimental [illegible]
A;Note: [illegible] the complete [illegible]
C;Superfamily: [illegible] xylanase A [illegible]
C;Keywords: [illegible]
F;[illegible] xylanase A [illegible] repeat [illegible]XA>


  Query Match             [illegible]
  Best Local Similarity   [illegible]; Pred. No. [illegible]
  Matches   [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible]

Qy          [illegible]NNQEGWK 1[illegible]

Db          [illegible]ENNEEGWK [illegible]


RESULT    [illegible]
[illegible]
[illegible] (KCM [illegible])
N;Alternate names: [illegible]
C;Species: [illegible] (KCM [illegible])
C;Date: [illegible] Mar [illegible] #sequence_revision [illegible] Mar [illegible] #text_change [illegible] Mar 2001
C;Accession: [illegible]
R;Hakamada, Y.; [illegible]; Kobayashi, T.; [illegible]; Kawai, S.; [illegible]; Ito, [illegible]
Biosci. Biotechnol. Biochem. [illegible]
A;Title: [illegible]
A;Reference number: [illegible]; MUID:[illegible]
A;Accession: [illegible]
A;Molecule type: DNA
A;Residues: [illegible]
A;Cross-references: [illegible]
A;Experimental [illegible] (KCM [illegible])
A;Accession: [illegible]
A;Molecule type: protein
A;Residues: [illegible]
C;Comment: [illegible] detergent
C;Genetics:
A;Gene: [illegible]
C;Keywords: [illegible]


  Query Match             [illegible]; Length 8[illegible]
  Best Local Similarity   [illegible]; Pred. No. [illegible]
  Matches   [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible]

Qy          [illegible]NNQEGWK [illegible]
                [illegible]
Db          [illegible]NNQEGWK [illegible]


RESULT    [illegible]
[illegible]
probable [illegible] - fission yeast (Schizosaccharo[illegible]
N;Alternate names: [illegible]
C;Species: [illegible]
C;Date: [illegible] Mar [illegible] #sequence_revision [illegible] Mar [illegible] #text_change [illegible]
C;Accession: [illegible]
R;[illegible] M.
[illegible] December [illegible]
A;Reference number: [illegible]
```

```
A;Accession: S62571
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-157 <PEA>
A;Cross references: EMBL:Z67961; NID:g1[illegible]; PIDN:CAA91899.1; PID:g10[illegible]
R;Pearson, D.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.
submitted to the EMBL Data Library, November 1995
A;Reference number: Z21801
A;Accession: T38586
A;Status: preliminary; translated from [illegible]
A;Molecule type: DNA
A;Residues: 1-157 <PE2>
A;Cross references: EMBL:Z67961; [illegible]
A;Experimental source: strain 972h-; cosmid [illegible]
R;Al-Khodairy, F.; Enoch, T.; Hagan, I.M.; Carr, A.M.
J. Cell Sci. 8, 475-486, 1995
A;Title: The S. pombe hus5 gene encodes a ubiquitin conjugating enzyme required for norm
A;Reference number: Z22481
A;Accession: T43397
A;Status: preliminary, translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-157 <ALK>
A;Cross-references: EMBL:X81846; PIDN:CAA[illegible]
enetics:
ene: HUS5; SPAC30D11.13
A;Map position: 1L
A;Introns: 22/3; 67/1; 75/3; 111/3; 1[illegible]/2
C;Function:
A;Description: required for efficient recovery from DNA damage or S phase arrest
C;Function: <CHRS>
A;Description: is possibly involved in chromosome segregation
C;Superfamily: human ubiquitin--protein ligase E2
C;Keywords: ligase


  Query Match             71.2%;  Score 4[illegible];  DB 2;  Length 157;
  Best Local Similarity   72.7%;  Pred. No. [illegible];
  Matches    8;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 LGILNNQEGWK 12
                | :|| :||||
Db         94 LSILNEEEGWK 104


RESULT    8
A29003
cellulase (EC 3.2.1.4), alkaline - Bacillus sp.
N;Alternate names: endo-1,4-beta-glucanase
C;Species: Bacillus sp.
C;Date: 21-May-1988 #sequence_revision 21-May-1988 #text_change 21-Jul-2[illegible]
C;Accession: A29003
kumori, F.; Kudo, T.; Narahashi, Y.; [illegible]
en. Microbiol. 132, 2329-2335, 1986
itle: Molecular cloning and nucleotide sequence of the alkaline cellulase gene from t
A;Reference number: A29003; MUID:8708544[illegible]
A;Accession: A29003
A;Molecule type: DNA
A;Residues: 1-800 <FUK>
A;Cross-references: GB:D00066; GB:N00[illegible]; NID:g216223; PIDN:BAA00045.1; PID:g216224
A;Experimental source: strain 1139
C;Function:
A;Description: hydrolysis of 1,4-beta-D-glucosidic linkages in beta-D-glucans such as ce
A;Pathway: cellulose degradation
C;Superfamily: Bacillus sp. alkaline cellulase; Thermotoga xylanase A amino-terminal rep
C;Keywords: glycosidase; hydrolase; polysaccharide degradation
F;585-726/Domain: Thermotoga xylanase A amino-terminal repeat homology <TXA>


  Query Match             65.2%;  Score 4[illegible];  DB 2;  Length 800;
  Best Local Similarity   77.8%;  Pred. No. 11;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 GILNNQEGW 11
```

```
                || ||:|
Db        200 GIPNNEEGW 208


RESULT    9
JT0611
cellulase (EC 3.2.1.4), alkaline - Bacillus sp. (strain KSM-64)
N;Alternate names: endo-1,4-beta-glucanase
C;Species: Bacillus sp.
C;Date: 09-Oct-1992 #sequence_revision 09-Oct-1992 #text_change 1[illegible]-Dec-1999
C;Accession: JT0611
R;Sumitomo, N.; Ozaki, K.; Kawai, S.; Ito, S.
Biosci. Biotechnol. Biochem. 56, 872-877, 1992
A;Title: Nucleotide sequence of the gene for an alkaline endoglucanase from an alkalo
A;Reference number: JT0611; MUID:92[illegible]
A;Accession: JT0611
A;Molecule type: DNA
A;Residues: 1-822 <SUM>
A;Cross references: GB:M84963; NID:g[illegible]; PIDN:AAA7[illegible]; PID:g[illegible]
C;Function:
A;Description: hydrolysis of 1,4-beta-D-glucosidic linkages in beta-D-glucans such as
A;Pathway: cellulose degradation
C;Superfamily: Bacillus sp. alkaline cellulase; Thermotoga xylanase A amino-terminal
C;Keywords: extracellular protein; glycosidase; hydrolase; polysaccharide degradation
F;585-726/Domain: Thermotoga xylanase A amino-terminal repeat homology <TXA>


  Query Match             65.2%;  Score 4[illegible];  DB 2;  Length 822;
  Best Local Similarity   77.8%;  Pred. No. 11;
  Matches    7;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 GILNNQEGW 11
                || ||:||
Db        200 GIPNNEEGW 208


RESULT   10
S42841
F16G12.1 protein - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 17-Mar-2000
C;Accession: S42841
R;Thomas, K.
submitted to the EMBL Data Library, February 1994
A;Reference number: S42837
A;Accession: S42841
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-885 <THO>
A;Cross references: EMBL:Z30317
C;Genetics:
A;Introns: 43/[illegible]
C;Superfamily: membrane alanyl aminopeptidase


  Query Match             62.1%;  Score 41;  DB 2;  Length 885;
  Best Local Similarity   70.0%;  Pred. No. 2[illegible];
  Matches    7;  Conservative    2;  Mismatches    [illegible];  Indels    0;  Gaps    0;

Qy          2 LGILNNQEGW 11
                || || ::|
Db        197 LGGLNNRKGW 206


RESULT   11
T30290
AAS surface protein - Staphylococcus saprophyticus
C;Species: Staphylococcus saprophyticus
C;Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 02-Sep-2000
C;Accession: T30290
R;Hell, W.; Meyer, H.G.W.; Gatermann, S.G.
Mol. Microbiol. 29, 871-881, 1998
A;Title: Cloning of aas, a gene encoding a Staphylococcus saprophyticus surface prote
```

```
A;Reference number: Z20[illegible]; MUID:[illegible]
A;Accession: [illegible]
A;Status: preliminary; translated from GB/EMBL/[illegible]
A;Molecule type: DNA
A;Residues: 1-14[illegible] <HEL
A;Cross-references: EMBL:AL[illegible]; NID:[illegible]
C;Genetics:
A;Gene: [illegible]

  Query Match             [illegible]
  Best Local Similarity   [illegible]
  Matches    6; [illegible]  Mismatches [illegible]

Qy          3 [illegible]ILNNQ[illegible]W [illegible]
              [illegible]
Db       1123 [illegible]IVNNK[illegible]W [illegible]

RESULT 12
F82663
hypothetical protein XF1[illegible] XF1[illegible] [illegible]
C;Species: Xylella fastidiosa
C;Date: 18 Aug 2000 #sequence_revision [illegible] Aug 2000 #text_change [illegible]
C;Accession: F82663; [illegible]82[illegible]
R;anonymous, The Xylella fastidiosa [illegible] for Nucleotide Sequen
Nature 406, 151-1[illegible] 2000
A;Title: The genome sequence of the plant pathogen [illegible]
A;Reference number: A8251[illegible]; MUID:2[illegible]
A;Note: for a complete list of authors see reference [illegible]
A;Accession: F82663
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-13[illegible] <[illegible]M>
A;Cross-references: GB:AE[illegible]; GB:AE[illegible]; NID:[illegible]; [illegible]AAF[illegible] [illegible] (No[illegible]
A;Experimental source: strain 9a5c
A;Accession: G82649
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-132 <S[illegible]>
A;Cross-references: GB:AE[illegible]; GB:AE[illegible]; NID:[illegible]; [illegible]AA[illegible] (N[illegible]
A;Experimental source: strain 9a5c
R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; [illegible] M.; Carrer, [illegible]
[illegible]Neto, E.; Docena, [illegible]; Ferry, [illegible]
[illegible]itted to GenBank, June 2000
[illegible]uthors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, [illegible]
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, [illegible]
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; [illegible] Marques, M.V.; Martins, [illegible]
A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, [illegible]
[illegible]; Nunes, L.R.; Oliveira, M.A.; de Oliveira, [illegible]
Rodrigues, V.; Rosa, A.J.; de M.; de [illegible]
A;Authors: da Silva, A.C.R.; da Silva, F.R.; [illegible] da Silveira, [illegible]
M.; Tsuhako, M.H.; Vallada, H.; Van Sluys, M.A.; [illegible]
A;Reference number: A[illegible]26
A;Contents: annotation
C;Genetics: GEN1
A;Gene: XF1598
C;Genetics: GEN2
A;Gene: XF1673

  Query Match             [illegible]
  Best Local Similarity   [illegible]
  Matches    [illegible]; Conservative [illegible]; Mismatches [illegible]

Qy          1 AI[illegible]ILNNQ[illegible]WK 12
              [illegible]
Db         81 AI[illegible]WMNARA[illegible]WK 92

RESULT 13
F81041
```

```
hypothetical protein NMA[illegible] [imported] - Neisseria meningitidis (strain Z2491 serogr
C;Species: Neisseria meningitidis
C;Date: [illegible] Mar 2000 #sequence_revision [illegible] Mar [illegible] #text_change [illegible]
C;Accession: F81041
R;Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, [illegible] M
[illegible] Mungall, K.L.; Quail, M.A.; [illegible]
Nature 4[illegible]
A;Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491
A;Reference number: A81[illegible]; MUID:[illegible]
A;Accession: F81041
A;Status: preliminary
A;Molecule type: DNA
A;Residues: [illegible]
A;Cross-references: GB:AL1[illegible]; GB:AL1[illegible]; NID:[illegible]4,4; PIDN:CAB[illegible]
A;Experimental source: serogroup A strain Z2491
C;Genetics:
A;Gene: NMA[illegible]
C;Superfamily: [illegible] hypothetical protein [illegible]

  Query Match             [illegible]
  Best Local Similarity   [illegible]; Pred. No. [illegible]
  Matches    [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible]

Qy          [illegible] A[illegible]NNQ[illegible]WK [illegible]
              [illegible]
Db          [illegible] A[illegible]IV[illegible]WK [illegible]

RESULT 14
T1[illegible]
NADH dehydrogenase chain [illegible] - Arabidopsis thaliana chloropla
C;Species: [illegible] Arabidopsis thaliana (mouse-ear cress)
C;Date: [illegible] #sequence_revision [illegible] #text_change [illegible]
C;Accession: [illegible]
R;[illegible] Mi[illegible] [illegible] A
M[illegible]
A;Title: [illegible] an example
A;Reference number: [illegible] MUID:[illegible]
A;Accession: [illegible]
A;Status: preliminary; translated from GB/EMBL/[illegible]
A;Molecule type: DNA
A;Residues: [illegible]
A;Cross-references: EMBL:A[illegible]; NID:[illegible]; PIDN:AA[illegible]
C;Genetics:
A;Genome: chloroplast
A;Note: [illegible]
C;Superfamily: NADH dehydrogenase [illegible]
C;Keywords: chloroplast; membrane; associated [illegible] NAD; oxidoreductase

  Query Match             [illegible]
  Best Local Similarity   [illegible]; Pred. No. [illegible]
  Matches    [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible]

Qy          [illegible] AI[illegible]NNQ[illegible]
              [illegible]
Db          [illegible] A[illegible]NQ[illegible]

RESULT 15
[illegible]
probable [illegible] (EC [illegible]) - winter squash
C;Species: [illegible] (winter squash)
C;Date: [illegible] #sequence_revision [illegible] #text_change [illegible] Aug [illegible]
C;Accession: [illegible]
R;[illegible] A
[illegible]
A;Title: [illegible] cDNA from pu
A;Reference number: [illegible] MUID:[illegible]
A;Accession: [illegible]
A;Status: preliminary; translated from GB/EMBL/[illegible]
A;Molecule type: mRNA
```

```
A;Residues: 1-346 <LAN>
A;Cross-references: EMBL:U63650; NID:q2289[illegible]; PID:q2289[illegible]
A;Experimental source: tissue_type endosperm; clone 83
C;Superfamily: 1-aminocyclopropane 1 carboxylate oxidase
C;Keywords: oxidoreductase


  Query Match             59.1%;  Score [illegible];  DB [illegible];  Length 346;
  Best Local Similarity   70.0%;  Pred. No. [illegible];
  Matches    7;  Conservative    1;  Mismatches    2;  Indels    [illegible];  Gaps    [illegible];

Qy          2 LGILNNQEGW 11
              | ||| :|||
Db        239 LQILNEKEGW 248


Search completed: August 22, 2001, 14:25:2
Job time: 79 sec
```

```
                          GenCore version 4.5
                 Copyright (c) 1993 - [illegible] Compugen Ltd.

OM protein - protein search, using sw model

Run on:          August 22, 2001, 14:24:[illegible]    Search time 112.[illegible] Seconds
                                                  (without alignments)
                                                  [illegible].688 Million cell updates/sec

Title:           US-08-659-254-2
Perfect score:   1429
Sequence:        1 MFLATFKLCAGSSYRHMRNM.....   ...QVLSQ[illegible]ANHLRKRLLA [illegible]7

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters:        [illegible]

[illegible]imum DB seq length: 0
[illegible]mum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       SPTREMBL_16:*
                1:  sp_archea:*
                2:  sp_bacteria:*
                3:  sp_fungi:*
                4:  sp_human:*
                5:  sp_invertebrate:*
                6:  sp_mammal:*
                7:  sp_mhc:*
                8:  sp_organelle:*
                9:  sp_phage:*
                10:  sp_plant:*
                11:  sp_rodent:*
                12:  sp_unclassified:*
                13:  sp_vertebrate:*
                14:  sp_virus:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                                SUMMARIES
               %
[illegible]t   Query
[illegible]o.  Score  Match Length DB  ID                          Description
-----------------------------------------------------------------------------
    1  1213.5   84.9    285   6  Q9GMD0                     q9gmd0 ovis aries
    2  1028.5   72.0    281  13  Q9DC09                     q9dc09 gallus gal
    3   977.5   68.4    289  13  Q9DC08                     [illegible]
    4     932   65.2    287  13  Q9DEB4                     q9deb4 oncorhynchu
    5     926   64.8    287  13  Q9DE07                     [illegible] salvelinus
    6   892.5   62.5    285  13  Q9DC10                     q9dc10 brachydani
    7     [illegible]   [illegible]    239  13  [illegible]     [illegible]
    8   211.5   14.8    580   5  Q9W145                     q9w145 drosophila
    9     195   13.6    219  11  Q9EJQ7                     [illegible] mus musculus
   10     165   11.5    180  11  Q9EPQ6                     [illegible] mus musculus
   11   134.5    9.4    478   5  Q19819                     q19819 caenorhabdi
   12   127.5    8.9    [illegible]   4  Q9NS[illegible]2   [illegible] homo sapien
   13   111.5    7.8   1047   5  Q9W3V4                     [illegible] drosophila
   14   103.5    7.2    995   4  Q9Y3M8                     q9y3m8 homo sapien
   15     101    7.[illegible]    29[illegible]   [illegible]  [illegible]   [illegible] caenorhabdi
   16    99.5    7.0    291  11  Q9JMD3                     q9jmd3 mus musculu
   17    [illegible]    7.0    359   4  [illegible]           [illegible] homo sapien
   18    97.5    6.8    243   2  Q9PB06                     q9pb06 xylella fas
   19      97    6.8    268   4  Q9UF[illegible]0           [illegible] homo sapien
   20    96.5    6.8   1281   2  Q9I2[illegible]            [illegible] pseudomonas
   21      96    6.7    322   5  O18[illegible]             [illegible] caenorhabdi
   22      94    6.6    686  10  Q9FV[illegible]            [illegible] arabidopsis
   23    93.5    6.5   1083   4  O14868                     o14868 homo sapien
   24    93.5    6.5   1091   4  O43[illegible]             [illegible] homo sapien
   25    92.5    6.5    296  11  Q9Z1X8                     q9z1x8 rattus norv
   26    92.5    6.5   3419  11  [illegible]                [illegible] rattus norv
   27    92.5    6.5  10223   2  Q54296                     q54296 streptomyce
   28      92    6.4    346   2  Q9K[illegible]             [illegible]
   29      92    6.4    483   2  Q9R[illegible]B8           [illegible] deinococcus
   30      92    6.4    663  10  Q9FH1[illegible]           [illegible] arabidopsis
   31      92    6.4   1434   5  Q23[illegible]             [illegible] drosophila
   32      92    6.4   1820   4  Q9P2P6                     q9p2p6 homo sapien
   33    91.5    6.4   4097   5  Q9V[illegible]             [illegible] drosophila
   34    90.5    6.3   1469  10  Q9XIE[illegible]           [illegible] arabidopsis
   35      90    6.3    176   5  Q9V1X3                     [illegible] drosophila
   36      90    6.3    720   5  Q9U9[illegible]5           [illegible] dictyosteli
   37      90    6.3   1247   5  Q9Y012                     q9y012 dictyosteli
   38      90    6.3   1247   5  Q9U911                     [illegible] dictyosteli
   39    89.5    6.3    356   2  Q9ZF78                     q9zf78 burkholderi
   40      89    6.2    598   4  Q9H[illegible]8            [illegible] homo sapien
   41      89    6.2   3663   5  Q9TW[illegible]            [illegible] caenorhabdi
   42      89    6.2   3674   5  Q9TY[illegible]            [illegible] caenorhabdi
   43    88.5    6.2    690   4  Q9HCB1                     [illegible] homo sapien
   44      88    6.2    679  10  O49417                     o49417 arabidopsis
   45    87.5    6.1    327  14  Q9YQY2                     q9yqy2 ranid herpe

                                ALIGNMENTS

RESULT    1
Q9GMD0
ID   Q9GMD0      PRELIMINARY;      PRT;    285 AA.
AC   Q9GMD0;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN.
GN   STAR.
OS   Ovis aries (Sheep).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Caprinae; Ovis.
OX   NCBI_TaxID=9940;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   West L.A., Horvat R.D., Roess D.A., Barisas B.G., [illegible] J.L.,
RA   Niswender G.D.;
RT   "Steroidogenic Acute Regulatory Protein and Peripheral-Type
RT   Benzodiazepine Receptor Associate at the Mitochondrial Membrane.";
RL   Submitted ([illegible] 2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF290[illegible]2; AAG02464.1; -.
SQ   SEQUENCE   285 AA;  [illegible] MW;  [illegible] CRC64;

  Query Match             84.9%;  Score 1213.5;  DB 6;  Length 285;
  Best Local Similarity   84.8%;  Pred. No. [illegible];
  Matches  235;  Conservative   22;  Mismatches   19;  Indels    1;  Gaps    1;

Qy          1 MFLATFK[illegible]CA[illegible]YRHMRNMKGLRH[illegible] 59
                [illegible]
Db          1 MLLATFK[illegible]CA[illegible]YRHVRSMKGLRQ[illegible]AAWIYQVRRRSS[illegible] 60

Qy         60 SQLEATLYSDQELSYIQQGEVAMQKAL[illegible]NNQ[illegible]WKKESQ[illegible]VLSKMVPDVGKVF 119
                [illegible]
Db         61 SQLEDSLYSDQELAYIQQGEEAMQKAL[illegible]WKKESKQ[illegible]VLSKVIPDVGKVF 12[illegible]

Qy        120 RLEVVVDQ[illegible]RLYEELVDRMEAM[illegible]KEIKVL[illegible]AAAA[illegible]NLVG 179
                [illegible]
Db        121 RLEVVV[illegible]RLYEELVERMEAM[illegible]KEIKVL[illegible]AAAA[illegible]NLVG 180

Qy        180 PRDFVSV[illegible]CKRRGSTCVLAGMAT[illegible]VLHPLA[illegible]SPSKTKL[illegible] 239
```

```
Db        181 PRDFVSVRC[illegible]EERG[illegible]VLAGIATI[illegible]YELMP[illegible]K[illegible]V[illegible]RAH[illegible] [illegible]

Qy        240 WELSIDEKGW[illegible]PKL[illegible]NQV[illegible]SQT[illegible]LANHLPEEL[illegible]A [illegible]

Db        241 WELSIDEKGW[illegible]PKL[illegible]NQV[illegible]LANHLPEE[illegible]


RESULT 2
Q9DG09
ID   Q9DG09     PRELIMINARY;     PRT;   [illegible] AA.
AC   Q9DG09;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN.
GN   STAR.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; [illegible]
OC   Archosauria; Aves; Neognathae; [illegible]
OC   Gallus.
OX   NCBI_TaxID=[illegible];
RN   [1]
RP   SEQUENCE FROM N.A.
RX   PubMed=11064[illegible];
RA   Bauer M.P., Bridgham [illegible], Langenau D.M., [illegible]
RT   "Conservation [illegible] steroidogenic [illegible] regulatory [illegible]
RT   structure and expression in vertebrates.";
RL   Mol. Cell. Endocrinol. [illegible]
DR   EMBL; AF220436; AAG2[illegible]
SQ   SEQUENCE   [illegible] AA;  [illegible] MW;  [illegible]

  Query Match           [illegible]
  Best Local Similarity [illegible]
  Matches  [illegible]; Conservative [illegible]; Mismatches [illegible]

Qy          1 MFLATFKL[illegible]AG[illegible]YRH[illegible]NMK[illegible]LRR[illegible]AVLAI[illegible]NWEA[illegible]

Db          1 MFPATFKL[illegible]AA[illegible]YQH[illegible]NV[illegible]LRR[illegible]AAVA[illegible]

Qy         [illegible] SG[illegible]EATIY[illegible]GY[illegible]EVAMQKA[illegible]NNL[illegible]WRE[illegible]

Db         [illegible] SR[illegible]EEKPE[illegible]MEMGY[illegible]

Qy        120 RLEVVVDQEMRRYEE[illegible]VEEMEAMGEWNENVK[illegible]

Db        121 RLEVVVDQ[illegible]LAVYSE[illegible]NMEQM[illegible]WNE[illegible]VK[illegible]VR[illegible]

Qy        180 PRDFVSVR[illegible]SSI[illegible]VLAGMAIHE[illegible]EMPR[illegible]V[illegible]

Db        181 PRDFVSVR[illegible]SSI[illegible]VLAGMAIHE[illegible]ANPR[illegible]

Qy        240 WELSIDEKGWE[illegible]EIN[illegible]VI[illegible]ANHLPEEL [illegible]

Db        241 WELSIDEKGWE[illegible]EIN[illegible]VI[illegible]QV[illegible]LANHLPEEM [illegible]


RESULT 3
Q9DG08
ID   Q9DG08     PRELIMINARY;     PRT;   [illegible] AA.
AC   Q9DG08;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN.
GN   STAR.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; [illegible]
OC   Amphibia; Batrachia; Anura; Mesobatrachia; [illegible]
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
```

```
RP   SEQUENCE FROM N.A.
RC   [illegible]
RX   [illegible]
RA   [illegible]
RT   [illegible] (StAR) protein [illegible]
RT   [illegible]
RL   [illegible]
DR   EMBL; [illegible]
DR   [illegible]
KW   [illegible]
SQ   SEQUENCE   [illegible] AA;  [illegible] MW;  [illegible]

  Query Match           [illegible]
  Best Local Similarity [illegible]
  Matches  [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible]

Qy          [illegible] ATFK[illegible]MENMK[illegible]AVLAI[illegible]NWEA[illegible]WMQVEERRSIL[illegible]

Db          [illegible] ATFK[illegible]ENM[illegible]AAVA[illegible]KWINQ[illegible]

Qy         [illegible] EATIY[illegible]VAMQKA[illegible]NNL[illegible]WRE[illegible]ENEEEV[illegible]KMVFQV[illegible]KVEE[illegible]

Db         [illegible] EEK[illegible]VEMGY[illegible]

Qy        [illegible] VVVD[illegible]VEEMEAMGEWNENVK[illegible]

Db        [illegible] AVVE[illegible]NVE[illegible]WNENVK[illegible]HEKAAL[illegible]

Qy        [illegible]

Db        [illegible]

Qy        [illegible]

Db        [illegible]


RESULT 4
[illegible]
ID   [illegible]     PRELIMINARY;     PRT;   [illegible] AA.
AC   [illegible]
DT   [illegible] MAR-[illegible] (TrEMBLrel. [illegible], Created)
DT   [illegible] MAR-[illegible] (TrEMBLrel. [illegible], Last sequence update)
DT   [illegible] MAR-[illegible] (TrEMBLrel. [illegible], Last annotation update)
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN.
GN   STAR.
OS   [illegible] (Rainbow trout) ([illegible]).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; [illegible]
OC   [illegible]
OX   NCBI_TaxID=[illegible];
RN   [1]
RP   SEQUENCE FROM N.A.
RC   [illegible]
RA   [illegible]
RT   "Molecular [illegible] (StAR) protein [illegible]
RT   cDNA [illegible]."
RL   Submitted ([illegible]) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; [illegible]
SQ   SEQUENCE   [illegible] AA;  [illegible] MW;  [illegible]

  Query Match           [illegible]
  Best Local Similarity [illegible]
  Matches  [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible]

Qy          [illegible] NWEA[illegible]WMQVEERRSIL[illegible]

Db          [illegible] MFLATF[illegible]MENM[illegible]KENAMVALDILNWEA[illegible]WMQVHVEEEIL[illegible]

Qy          [illegible] VAMQKA[illegible]NNL[illegible]WRE[illegible]ENEEEV[illegible]KMVFQV[illegible]KVE[illegible]
```

```
Db         60 RIEEEQGYNFAFVSYVF[illegible]EFAL[illegible]ESISI[illegible] 11[illegible]

Qy        120 PLEVVVPQPMDPLYEELVDRMEAMCEWNPNVKE[illegible]KV[illegible]Q[illegible]IGKETVI[illegible]HEE[illegible]AAAAMNLVG 17[illegible]
              :|||::||   | || |||   || ||:|||  ||||:|:  ::||::|::|:|::          |:::
Db        120 KLEVLLPQRSDNLYVELVCNMEQMCDWNPNVKE[illegible]KI[illegible]ETMVIHE[illegible]NVV[illegible] 17[illegible]

Qy        180 PRDFVSVRCTKRRGSTCVLACMATHPCEM[illegible]EQ[illegible]VIRAEHCPTCMVLHPLAG[illegible]KIKLI [illegible]
              ||||:||||| |||||| ||||:|        | |   ::|::|||||:|::   :   :: |
Db        180 PRDFVSVRCAKRRG[illegible]LAGMSTQHPIM[illegible]NKIKEI [illegible]

Qy        240 WLLSIDLKGWLPKTIINQVLSQTQIEFANHLRK[illegible] 274
              ||||||||||:| ||||:||||||::||||     :
Db        240 WLLSIDLKGWIPKTIINKVLSQTQVEFANH[illegible] 27[illegible]


RESULT    5
Q9DE06
ID   Q9DE06        PRELIMINARY;       PRT;   287 AA.
AC   Q9DE06;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
     STEROIDOGENIC ACUTE REGULATORY PROTEIN.
     STAR.
     Salvelinus fontinalis (Brook trout) (Brook char).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC   Protacanthopterygii; Salmoniformes; Salmonidae; Salvelinus.
OX   NCBI_TaxID 8038;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=BRAIN, TESTIS, HEAD, KIDNEY, SKIN, AND LIVER;
RA   Goetz F.W.;
RT   "Characterization of the brook trout [illegible] and MLN64 homologs.";
RL   [illegible]
DR   EMBL; AF232215; AAG39689.1; -.
SQ   SEQUENCE   287 AA;  31884 MW;  291F41[illegible]E9B[illegible]C64;

  Query Match             64.8%;  Score [illegible];  DB 13;  Length 287;
  Best Local Similarity   63.3%;  Pred. No. [illegible]e-68;
  Matches  174;  Conservative   45;  Mismatches   54;  Indels   [illegible];  Gaps   [illegible];

Qy          1 MFLATFKLCAGSSYRHMRNMKGLRHQAV[illegible]AI[illegible]NWLAI[illegible] [illegible]
              |  ||||||||| |||| ||| |||   |: ||  | |   | |  :      :   | | | :
Db          1 MLPATFKLCAGISYRHIRNMKGLRKNAM[illegible]NMLA[illegible]GPNI[illegible] [illegible]

Qy         63 QLEATL[illegible]YSDQELSYI[illegible]QGEVAMQKALGILNN[illegible]WKKESQQENGDEVLSKMVPEVGKVF 11[illegible]
              ::|     |::  ::||::||| |:|||::  :|  :  :  |          |:|:|||:: |:
Db         60 FIPEEQGYNFAFVSYVFQGHEALQEIIS[illegible]KVF 11[illegible]

          120 PLEVVVPQPMDPLYEELVDRMEAMGEWN[illegible]NVKE[illegible]KVI[illegible]GKETVIIHEEAAAAAMNLV[illegible] 1[illegible]
              :|||::||   | || |||   || ||:|||||  : :|::|::|::|||::      |:::
Db        120 KLEVLLPQRSDNLYCELVCNMEQMGDWN[illegible]NV[illegible]EI[illegible]EIMVI[illegible]EVSGP[illegible]NVV[illegible] 1[illegible]

Qy        180 PRDFVSVRCTKRRGSTCVLACMATHPCEMP[illegible]VIRA[illegible]HCPTCMVLHPLAG[illegible]KIKLI 2[illegible]
              |||||||| |||||| ||||:|       |||| |::|||:||||||:|: | |  :::|:  |
Db        180 PRDFVSVRCAKRRGSTCFLAGMSTQHP[illegible]NVVRAENG[illegible]CIVML[illegible]NKIKE[illegible] 2[illegible]

Qy        240 WLLSIDLKGWLPKTIINQVLSQTQIEFANHLRK[illegible] 274
              |||||||||:||||||:||||||::|||| :
Db        240 WLLSIDLKGWIPKTIINKVLSQTQVDFANHLR[illegible]M 274


RESULT    6
Q9DG10
ID   Q9DG10        PRELIMINARY;       PRT;   [illegible] AA.
AC   Q9DG10;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN
GN   STAR.
```

```
OS   Brachydanio rerio (Zebrafish) (Zebra danio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC   Cypriniformes; Cyprinidae; Rasborinae; Danio.
OX   NCBI_TaxID 7955;
RN   [1]
RP   SEQUENCE FROM N.A
RX   PubMed 11064158;
RA   Bauer M.P., Bristham J.T., Landon [illegible] M., [illegible] Goetz F.W.;
RT   "Conservation of steroidogenic acute regulatory (StAR) protein
RT   structure and expression in vertebrates.";
RL   Mol. Cell. Endocrinol. 168:119-125(2000).
DR   EMBL; AF22[illegible]; AAG28593.1; -.
SQ   SEQUENCE   285 AA;  31759 MW;  [illegible]64;

  Query Match             62.5%;  Score 8[illegible];  DB 13;  Length 285;
  Best Local Similarity   59.1%;  Pred. No. [illegible];
  Matches  162;  Conservative   52;  Mismatches   59;  Indels    1;  Gaps    1;

Qy          1 MFLA[illegible]KLCAG[illegible]YRHMRNMKGLRHQAV[illegible]NWLAI[illegible]DSSL[illegible]WMGQVRRRSSLLGS 6[illegible]
              |  |||||||| ||||||||| |||      :    :   :     :    :  :   : :||    :
Db          1 MLPATFKLCAGISYRHMRNMKGLRKNAM[illegible]PGA[illegible]NHHRRRSSLLSS 5[illegible]

Qy         6[illegible] QLEATL[illegible]QGEVAMQKALG[illegible]WKK[illegible]QQENG[illegible]VLSKMVP[illegible]VGKV[illegible] 1[illegible]
                :    ||: :   |:|||: |:||:: ||  :  |:  |  :  |  :  |:||::| :||||:
Db         60 FIAEEY[illegible]EAL[illegible]YVQ[illegible]EALQKSIG[illegible]NE[illegible]VMGKVIE[illegible]GKVFK 11[illegible]

Qy        121 LEVVVPQPMDPLYEELVDRMEAMGEWNPNVKE[illegible]KVI[illegible]GKETV[illegible]HEEAAAAMNLVG 1[illegible]
              |||  ::|       |:||||  ||  ||||||:       |   ||  :|:  ||:|:|      |::::|
Db        120 LEVTLPQ[illegible]YEELVCNMEQMGEWNPNV[illegible]VK[illegible]NVV[illegible]H 1[illegible]

Qy        181 RDFVSVRCTKRRGSTCVLACMATHPC[illegible]ME[illegible]GV[illegible]AEHCPTCMVLHPLAGTSKIKLIW 24[illegible]
              ||||::|   :|:|:|| ||||:|      |||        :  :|   :       :   :  |:   |
Db        1[illegible] R[illegible]FVNVRGAK[illegible]LA[illegible]NEIKLIW 2[illegible]

Qy        241 LLSIDLKGWLPKTIINQVLSQTQIEFANHLRK[illegible] 274
              ||||:||     ::::|:||:||||::|  ||    :
Db        240 LLSLDLKGWIPKTVINKVLSQTQVDFVNH[illegible]M [illegible]


RESULT    7
Q9DFS4
ID   Q9DFS4        PRELIMINARY;       PRT;   2[illegible] AA.
AC   Q9DFS4;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   MLN64 LIKE PROTEIN (FRAGMENT).
GN   MLN64.
OS   Brachydanio rerio (Zebrafish) (Zebra danio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC   Cypriniformes; Cyprinidae; Rasborinae; Danio.
OX   NCBI_TaxID 7955;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   PubMed 11064158;
RA   Bauer M.P., Bristham J.T., Landon [illegible] M., [illegible] A.L., Goetz F.W.;
RT   "Conservation of steroidogenic acute regulatory (StAR) protein
RT   structure and expression in vertebrates.";
RL   Mol. Cell. Endocrinol. 168:119-125(2000).
DR   EMBL; AF2[illegible]; AAG28[illegible].1; -.
FT   NON_TER       1       1
SQ   SEQUENCE   2[illegible] AA;  27345 MW;  [illegible]A[illegible]4;

  Query Match             27.4%;  Score [illegible];  DB 13;  Length 2[illegible];
  Best Local Similarity   36.8%;  Pred. No. [illegible]e-14;
  Matches   82;  Conservative   49;  Mismatches   92;  Indels   [illegible];  Gaps   [illegible];

Qy         52 RRRSSLLGSQLEATLYSDQELSYIQQGEVAMQKAL[illegible]NNQEGWKKESQQENGDEVLSKM [illegible]
              | |  |              :::|| :::::           :     :    |  :   |
Db        9 EVEDDLEFE[illegible]RA[illegible]KAEV[illegible]AMAVVE[illegible]NWE[illegible]

Qy      112 VE[illegible]VGKV[illegible]VVVL[illegible]M[illegible]YE[illegible]MEAM[illegible]WN[illegible]NV[illegible]AA
                  ||
Db       69 IEEHGKIE[illegible]AEIVYQEV[illegible]EEKMV[illegible]

Qy      172 AAA[illegible]NLV[illegible]VSV[illegible]TKEE[illegible]VLA[illegible]MAIHE[illegible]MEE[illegible]
                        |    | :|
Db      129 GAA[illegible]GVV[illegible]VNV[illegible]VEKEK[illegible]MAINEN[illegible]

Qy      232 [illegible]SKTKLEWE[illegible]WEEK[illegible]NLV[illegible]ANE[illegible]
              ||       |     ||
Db      189 NE[illegible]VCIE[illegible]WV[illegible]AA[illegible]ME[illegible]


RESULT 8
Q9W145
ID   Q9W145      PRELIMINARY;   [illegible]   AA.
AC   Q9W145;
DT   01 MAY 2000 (TrEMBLrel. [illegible] Created)
DT   01 MAY 2000 (TrEMBLrel. [illegible] Last sequence update)
DT   01 MAR 2001 (TrEMBLrel. [illegible] Last annotation update)
DE   CG3522 PROTEIN.
GN   CG3522.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; [illegible]
OC   Pterygota; Neoptera; Endopterygota; Diptera; [illegible]
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RX   MEDLINE=[illegible]; PubMed=[illegible];
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., [illegible]
RA   Amanatides P.G., Scherer S.E., Li P.W., [illegible]
RA   George R.A., Lewis S.E., Richards S., Ashburner M., [illegible]
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., [illegible]
RA   Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., [illegible]
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., [illegible]
RA   Abril J.F., Agbayani A., An H.J., Andrews [illegible]
RA   Ballew R.M., Basu A., Baxendale J., [illegible]
RA   Beeson K.Y., Benos P.V., Berman B.P., [illegible]
RA   Borkova D., Botchan M.R., Bouck J., [illegible]
RA   Burtis K.C., Busam D.A., Butler H., [illegible]
RA   Cherry J.M., Cawley S., Dahlke C., [illegible]
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., [illegible]
RA   Dodson K., Doup L.E., Downes M., [illegible]
RA   Durbin K.J., Evangelista C.C., [illegible]
RA   Foster C., Gabrielian A.E., Garg N.S., [illegible]
RA   Glodek A., [illegible]
RA   Harris N.L., Harvey D., [illegible]
RA   Hoskins R.A., Houck J., [illegible]
RA   Ibegwam C., Jalali M., Kalush F., [illegible]
RA   Kimmel B.E., Kodira C.D., Kraft C., [illegible]
RA   Lasko P., Lei Y., Levitsky A.A., [illegible]
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., [illegible]
RA   Merkulov G., Milshina N.V., Mobarry C., [illegible]
RA   Mount S.M., Moy M., Murphy B., Murphy L., [illegible]
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., [illegible]
RA   Palazzolo M., Pittman G.S., Pan S., [illegible]
RA   Reinert K., Remington K., Saunders R.D., [illegible]
RA   Shue B.C., Siden-Kiamos I., [illegible]
RA   Spier E., Spradling A.C., Stapleton M., Strong R., [illegible]
RA   Svirskas R., Tector C., Turner R., Venter E., [illegible]
RA   Wang Z.Y., Wassarman D.A., [illegible]
RA   Williams S.M., Woodage T., Worley K.C., [illegible]
RA   Ye J., Yeh R.F., [illegible]
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., [illegible]
RA   Gibbs R.A., Myers E.W., [illegible]
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
DR   EMBL; AE[illegible]; AAF[illegible]
DR   HSSP; Q14049; [illegible]
```

```
[illegible]
Query Match [illegible] Score [illegible] DB [illegible] Length [illegible]
Best Local Similarity [illegible] Pred. No. [illegible]
Matches [illegible] Conservative [illegible] Mismatches [illegible] Indels [illegible] Gaps [illegible]

Qy      [illegible] VAMLKA[illegible]NN[illegible]WEE[illegible]ENE[illegible]VI[illegible]EMV[illegible]
Db      [illegible] WEV[illegible]VN[illegible]

Qy      [illegible]VVVV[illegible]MEAM[illegible]WN[illegible]NVE[illegible]AAAAA[illegible]
Db      [illegible]AE[illegible]EKWNLA[illegible]YV[illegible]

Qy      [illegible]

Db      [illegible]IAQ[illegible]

Qy      [illegible]

Db      [illegible]LAKEK[illegible]

Qy      [illegible] VLA[illegible]MAIHE[illegible]EME[illegible]V[illegible]VIHE[illegible]

Db      [illegible]LANNMIKENV[illegible]VGKIEE[illegible]VWVESAV[illegible]YAAVE[illegible]NIV[illegible]EAEEEIV[illegible] 24

Qy      [illegible] SK[illegible]KL WEE[illegible]WEEK[illegible]NLV[illegible]EEANE[illegible] [illegible]4

Db      [illegible]WVE[illegible]AA[illegible]SMI[illegible]


RESULT [illegible]
[illegible]
ID   [illegible]      PRELIMINARY;   [illegible]
AC   [illegible]
DT   [illegible] MAR [illegible] (TrEMBLrel. [illegible] Created)
DT   [illegible] MAR [illegible] (TrEMBLrel. [illegible] Last sequence update)
DT   [illegible] MAR [illegible] (TrEMBLrel. [illegible] Last annotation update)
DE   [illegible]
GN   [illegible]
OS   Mus musculus (Mouse).
OC   [illegible]
OC   Mammalia; Eutheria; Rodentia; [illegible]
OX   NCBI_TaxID=[illegible]
RN   [illegible]
RP   SEQUENCE FROM N.A.
RC   [illegible]
RA   [illegible]
RA   [illegible]
RT   [illegible]
RT   [illegible]
RL   [illegible]
DR   EMBL; [illegible]; AAF[illegible]
DR   [illegible]
SQ   SEQUENCE [illegible] AA; [illegible] MW; [illegible]

Query Match [illegible]
Best Local Similarity [illegible] Pred. No. [illegible]
Matches [illegible] Conservative [illegible] Mismatches [illegible] Indels [illegible] Gaps [illegible]

Qy      [illegible]EVAM[illegible]KA[illegible]NQ[illegible]WEKE[illegible]N[illegible]EMV[illegible]VVV[illegible]

Db      [illegible]EAV[illegible]NV[illegible]WE[illegible]NIY[illegible]

Qy      [illegible]MEAM[illegible]WNENVEE[illegible]AAAAA[illegible]

Db      [illegible]AAMK[illegible]
```

```
Qy     187 FCTKKRGSICVLAGMAIHGE  MPEQSCVIRA[illegible]PMVLHH[illegible] 244
                |:      ::  |||       | : | :    : |       || | :|
Db     126 VLVKKY [illegible]SNATRV[illegible]V[illegible]PCEPL[illegible]NKINLVI[illegible] [illegible]

Qy     245 DLFCWI[illegible]NQV[illegible]Q[illegible]FANHLKK 272
           || |:||::::     ::  ||  :|:|
Db     185 DLSGYLPQSVVDSFFPRSMAEFYPNLQK 212


RESULT 10
Q9EPQ6
ID   Q9EPQ6      PRELIMINARY;      PRT;   [illegible] AA.
AC   Q9EPQ6;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   STAR RELATED PROTEIN P3-15E/P3-16E (FRAGMENT).
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID 10090;
     [1]
     SEQUENCE FROM N.A.
     STRAIN=SWISS WEBSTER/NIH;
RA   Wines M.E., Lee L., Katari M.S., Zhang [illegible], DeRossi C., Shi Y.,
RA   Perkins S., Feldman M., McCombie W.R., [illegible]
RT   "Identification of mesoderm development (mesd) candidate genes by
RT   comparative mapping and genome sequence analysis.";
RL   Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY007809; AAG41057.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   180 AA;  20099 MW;  AA[illegible] CRC64;

  Query Match             11.5%;  Score [illegible];  DB 11;  Length 180;
  Best Local Similarity   26.7%;  Pred. No. [illegible];
  Matches   44;  Conservative   35;  Mismatches   72;  Indels   14;  Gaps    5;

Qy     116 GKVFELF VVVIQEMLFLYFFLVTEMEAMG      [illegible]PIVIIHEL 169
              | ::| | ::    |       ||: | ::          :|: ||   :::| |     ::
Db      15 GNLYRGEGILCGTP-----EEVWDCIKPVAS[illegible]KWHDNVSSFE[illegible]VSFIS [illegible]

Qy     170 AAAAAGNLVGPRDFVSVRCIKRRGSICVLACMAI[illegible]MPEQSCVIRAEHG[illegible]MVLH 2[illegible]
               :||  |: |||||  :    |:          ::    |,        | : | :|  : ,
Db      70 TPSAAMKLISPRDFV[illegible]VKY [illegible]VF[illegible]GENH[illegible] 1[illegible]

Qy     229 PLAGSPSKIKLTWLLS[illegible]LKGW[illegible]NQV [illegible]ANHLKK 272
               || | |:|| |          || |:||::::       :      |  :,:
Db     129 PLPGDPNKTNLVTPPQTDLSGYLPQSVV[illegible]SMAEFYPNLQK 173


[illegible]LT 11
Q19819
ID   Q19819      PRELIMINARY;      PRT;   4[illegible] AA.
AC   Q19819;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   COSMID F26F4.
GN   F26F4.4.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID 6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RX   MEDLINE 94150718; PubMed 7906398;
RA   Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA   Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA   [illegible] M., [illegible]
RA   Gardner A., Green P., Hawkins T., [illegible] L., Jier M., Johnston L.,
RA   Jones M., Kershaw J., Kirsten J., [illegible] N., La[illegible] P.,
RA   Lightning J., Lloyd C., Mcmurray A., M[illegible] B., [illegible] J.,
RA   Parsons J., Percy C., Ritken L., [illegible] A., [illegible] R.,
RA   Smaldon N., Smith A., Sonnhammer E., Staden R., Sulston J.,
RA   Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA   Watson A., W[illegible] L., Wilkinson Sproat J., Wohldman P.;
RT   "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT   elegans.";
RL   Nature 368:32-38(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   Fulton L.;
RL   Submitted (NOV-1994) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   Waterston R.;
RL   Submitted (MAR-1995) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U12964; AAA91219.1; -.
DR   InterPro; IPR002913; -.
DR   Pfam; PF01852; START; 1.
DR   SMART; SM00234; START; 1.
SQ   SEQUENCE   478 AA;  54924 MW;  [illegible] CRC64;

  Query Match             9.4%;  Score [illegible];  DB 5;  Length 478;
  Best Local Similarity   21.[illegible]%;  Pred. No. [illegible];
  Matches   5[illegible];  Conservative   53;  Mismatches   98;  Indels   41;  Gaps   11;

Qy      53 KRSGLLKGGLEDAILYGDQHLGYIQQGLVAM[illegible]KMV 112
              ||     |:  ||         :|                 : :  : :  :  ::  :  :    | ::
Db     213 KRQQNDGNQ[illegible]AVPGRVS--   -    [illegible]ARSRL[illegible] 2[illegible]

Qy     113    PDVGKVLRL VVVIQEMDRLYHELVLLP[illegible]       NENVK         [illegible]K 154
              ||   : |   |: :  :  :|:: ||:: |    :| |        | |           ,  ,
Db     293 VPPLTKRLL  [illegible]LRDCLDQV-HELM [illegible]ANL [illegible]PENYFLV[illegible]K 34[illegible]

Qy     155 VLQKIGKDFVIIHELAAAAAGNLVGPRDFV[illegible]VRCIKRR[illegible]CVLA[illegible] MAIP[illegible]QS 21[illegible]
           :: :    |  :  :  ::| |     :  ,   :   |          |  |        ::
Db     349 MIAHLDNATD[illegible]YYSVSAPAMRGYISSKDE[illegible]KKIK [illegible]ADV[illegible]GYIVGVEGNL[illegible]NNG 4[illegible]

Qy     213 G--VIRAEHGP[illegible]CMVLHPLAGSPSKIKL[illegible]W[illegible]NQVLSQ[illegible]ANHL 270
                ::|   : |:   ::   |          |     ,  ,      |  :  :   :        |: |:
Db     409 NPKIVRGHNP[illegible]G  MIKILKEDAGVIY[illegible]MK[illegible]LFK[illegible]VHM[illegible]YT[illegible]HV 4[illegible]

Qy     271 RKRLE 27[illegible]
            ::  |
Db     464 KRMNE 46[illegible]


RESULT 12
Q9NSY2
ID   Q9NSY2      PRELIMINARY;      PRT;   [illegible] AA.
AC   Q9NSY2;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   HYPOTHETICAL [illegible]2.1 KDA PROTEIN.
GN   DKFZP434G2416.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID 9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=TESTIS;
RA   Koehrer K., Beyer A., Mewes H.W., [illegible] J., Wiemann S.;
RL   Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL13[illegible]; CAB70862.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   [illegible] AA;  12[illegible] MW;  [illegible]B[illegible] CRC64;
```

```
  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches [illegible]

Qy    177 [illegible]VGPRDEVSVE[illegible]TKR[illegible]VLAGMA[illegible]
              |||        |:
Db      3 [illegible]GPRDEVG[illegible]VKE[illegible]SSNA[illegible]HVIDH[illegible]

Qy    235 K[illegible]K[illegible]TW[illegible]KGW[illegible]K[illegible]NQV[illegible]EANL[illegible]
              |  |            ||
Db     62 K[illegible]N[illegible]VI[illegible]GY [illegible]NVV[illegible]M[illegible]AN[illegible]


RESULT 13
Q9W3V4
[illegible]  Q9W3V4     [illegible]  MIN  Y  [illegible]  AA.
[illegible]  Q9W3V4;
[illegible]  01-MAY-2000 (TrEMBLrel. [illegible] Created)
DT    01-MAY-2000 (TrEMBLrel. [illegible] Last sequence update)
DT    01-OCT-2000 (TrEMBLrel. [illegible] Last annotation update)
DE    CG4559 PROTEIN.
GN    CG4559.
OS    Drosophila melanogaster (Fruit fly).
OC    Eukaryota; Metazoa; Arthropoda; [illegible]
OC    Pterygota; Neoptera; Endopterygota; Diptera; [illegible]
OC    Ephydroidea; Drosophilidae; Drosophila.
OX    NCBI_TaxID=7227;
RN    [1]
RP    SEQUENCE FROM N.A.
RC    STRAIN=BERKELEY;
RX    MEDLINE=20196006; PubMed=[illegible];
RA    Adams M.D., Celniker S.E., Holt R.A., [illegible]
RA    Amanatides P.G., Scherer S.E., [illegible] W[illegible]
RA    George R.A., Lewis S.E., Richards S., Ashburner M., [illegible]
RA    Sutton G.G., Wortman J.R., Yandell M.D., [illegible]
RA    Brandon R.C., Rogers Y.-H.C., [illegible]
RA    Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., [illegible]
RA    Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., [illegible]
RA    Ballew R.M., Basu A., Baxendale J., [illegible]
RA    Beeson K.Y., Benos P.V., Berman B.P., [illegible]
RA    Borkova D., Botchan M.R., Bouck J., [illegible]
RA    Burtis K.C., Busam D.A., Butler H., [illegible]
RA    Cherry J.M., Cawley S., [illegible]
RA    de Pablos B., [illegible]
[illegible]  Dodson K., [illegible] Dewar M[illegible]
[illegible]  Durbin K.J., [illegible]
RA    Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., [illegible]
RA    Glodek A., [illegible]
RA    Harris N.L., [illegible]
RA    Hostin D., [illegible] K[illegible]
RA    Ibidi M., K[illegible]
RA    Kimmel B.E., K[illegible]
RA    Lasko P., Lei Y., [illegible] A.A.
RA    Liu X., Mattei B., M[illegible]
RA    Merkulov G., [illegible]
RA    Mount S.M., M[illegible]
RA    Nelson D.R., Nelson K.A., Nixon K., N[illegible]
RA    Pallfreyman M., [illegible]
RA    Reinert K., [illegible]
RA    Shue B.C., Siden-Kiamos I., [illegible]
RA    Spier E., Spradling A.C., [illegible]
RA    Strinkas R., [illegible] Turner R., Venter [illegible]
RA    Wang Z.-Y., W[illegible]
RA    Williams S.M., W[illegible]
RA    Yeh R.[illegible]
RA    Zhang X.H., Zhong F.N., [illegible]
RA    Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT    "The genome sequence of Drosophila melanogaster.";
RL    Science 287:2185-2195(2000).
DR    EMBL; AE003484; AAF5[illegible].
DR    FlyBase; FBgn[illegible]
DR    InterPro; IPR[illegible]
```

```
[illegible]

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches [illegible]

Qy    [illegible]

Db    [illegible]

Qy    [illegible]EVAMLKA[illegible]NNG[illegible]WKK[illegible]

Db    [illegible]

Qy    [illegible]

Db    [illegible]

Qy    [illegible]AAAAGN[illegible]

Db    [illegible]

Qy    [illegible]EAN[illegible]

Db    [illegible]

Qy    [illegible]

Db    [illegible]


RESULT 1[illegible]
[illegible]
[illegible]  PRELIMINARY [illegible]
[illegible]
[illegible]

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches [illegible]

Qy    [illegible]

Db    [illegible]

Qy    [illegible]

Db    [illegible]
```

```
Qy     166 THELAAAAACNLVGP---RDFVSVRCIKR [illegible] IVLAGMATHHGEMPEQSGVIRAEH 219
                :      | : |   ||||:|  |     :      '  ::    !  :  : : :
Db     879 YQYVL----NSMAPHPSKDFVVLRTWKTELPE [illegible] EHAQ [illegible] VIAVV [illegible]

Qy     220 GPTCMVIHPLAGSPSKIKLIWLLSIDLKGWLPK 2 2
             :  :: |  ||  |::|| :  |||||  ,
Db     932 MDSQYLIEP GGS-GKSRLIH CKIDLKGHSPE [illegible]


RESULT  15
Q21130
ID   Q21130      PRELIMINARY;    PRT;   296 AA.
AC   Q21130;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01 NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01 MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   K02D3.2 PROTEIN.
GN   K02D3.2.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis
OX   NCBI_TaxID 6239;
     [1]
     SEQUENCE FROM N.A.
     Bardill S.;
RL   Submitted (MAR 1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE-94150718; PubMed-7906398;
RA   Wilson R., Ainscough R., Anderson K , [illegible] C., [illegible] M.
RA   Bonfield J., Burton J., Connell M., [illegible] T., Cooper J., [illegible] A.,
RA   Craxton M., Dear S., Du Z., Durbin R., [illegible] A., Fulton [illegible]
RA   Gardner A., Green P., Hawkins T., [illegible] L., Jier M., [illegible] L.,
RA   Jones M., Kershaw J., Kirsten J., [illegible] N., Latreille P.,
RA   Lightning J., Lloyd C., Mcmurray A., Mortimore B., [illegible] M.,
RA   Parsons J., Percy C., Rifken L., [illegible] A., Saunders D., [illegible] R.,
RA   Smaldon N., Smith A., Sonnhammer E., [illegible] R., Sulston J.,
RA   Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA   Watson A., Weinstock L., Wilkinson-Sproat J., Wohldman P.;
RT   "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT   elegans.";
RL   Nature 368:32-38(1994).
DR   EMBL; Z70209; CAA94145.1; .
DR   InterPro; IPR002913; .
DR   Pfam; PF01852; START; 1.
DR   SMART; SM00234; START; 1.
SQ   SEQUENCE   296 AA;  33385 MW;  02BC17E3 [illegible] CRC64;


  Query Match             7.1%;  Score [illegible];  DB 5;  Length 296;
  Best Local Similarity   27.6%;  Pred. No. [illegible];
  Matches   37;  Conservative   27;  Mismatches   58;  Indels   12;  Gaps    7;

Qy     145 FWNPNVFEIKVLQRIGFET-VITHELAAAAACNLV [illegible] EVSVR [illegible] VLAG 200
            :|:        ::: |  || || |:            |  ||    |              :   :
Db     144 KWDSQSACTSLVEEISYPTSVILHKTKCRMMC [illegible] ILLD [illegible] 2 [illegible]

Qy     201 MATHHGEMPEQSGVIRAEHGPTCMVLHPLA [illegible] SKIKLIWLLSID- LKGWLPK [illegible] IN V [illegible]
               |  :  |   :||:||:  ||  ::|:|,          |:|  ::   |   |:|  :::  ::  :
Db     203 SVDH-EKSPRIAGVVRAQIYPTLLMLNP [illegible] VK [illegible] SIVQA [illegible] VP [illegible] 2 [illegible]

Qy     259 LSQIQIEFANHLRK 272
            :  :      :    |:|
Db     258 MPKSIASFFEDLQK 271


Search completed: August 22, 2001, 14:27: [illegible]
Job time: 194 sec
```

```
                         GenCore version [illegible]
              Copyright (c) 1993 - [illegible] [illegible] Ltd.


OM protein - protein search, using sw model

Run on:          August 22, 2001, 14:30:0[illegible] ; Search time [illegible] Seconds
                                                (without alignments)
                                                5[illegible]4 Million cell updates/sec

Title:           US-08-659-254-3
Perfect score:   80
Sequence:        1 AEHGPTCMVLHPLA 14

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        19733[illegible] seqs, 20590346 residues

Total number of hits satisfying chosen parameters:      197[illegible]

[illegible]mum DB seq length: 0
[illegible]mum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       Issued_Patents_AA:*
                1: [illegible] pep:*
                2: [illegible] pep:*
                3: [illegible] pep:*
                4: [illegible] pep:*
                5: [illegible] pep:*
                6: [illegible] pep:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                    SUMMARIES
                     %
Result              Query
  No.     Score     Match Length DB  ID                          Description
---------------------------------------------------------------------------------
    1        80     100.0     14  2  US-08-[illegible]           [illegible]
    2        80     100.0     14  4  US-08-659-254-3             Sequence 3, Appli
    3        80     100.0    276  4  US-08-659-254-2             Sequence 2, Appli
    4        80     100.0    284  2  US-08-5[illegible]-2        Sequence 2, Appli
    5        80     100.0    285  1  US-08-41[illegible]-140-2   Sequence 2, Appli
    6        80     100.0    285  4  US-08-659-254-18            Sequence 18, Appli
    7        53      66.2      9  2  US-08-[illegible]           Sequence 7, Appli
    8        53      66.2      9  4  US-08-659-254-7             Sequence 7, Appli
    9        38      47.5   1285  2  US-08-54[illegible]-6       Sequence 6, Appli
   10        38      47.5   1285  3  US-08-65[illegible]-6       Sequence 6, Appli
   11        38      47.5   1285  4  US-08-951[illegible]-6      Sequence 6, Appli
   12        38      47.5   1285  5  PCT-US95[illegible]         Sequence 6, Appli
   13        38      47.5   1286  4  US-09-26[illegible]-140-3   Sequence 3, Appli
   14        38      47.5   1299  4  US-08-46[illegible]C-62     Sequence 62, Appli
   15        37      46.2    215  1  US-08-45[illegible]A-1      Sequence 1, Appli
   16        37      46.2    716  2  US-08-7[illegible]-082-1    Sequence 1, Appli
   17        37      46.2    716  4  US-09-2[illegible]-219-1    Sequence 1, Appli
   18        37      46.2   3080  6  522342[illegible]-4         Patent No. 522[illegible]
   19        35      43.8     31  2  US-08-16[illegible]-48B-2   Sequence 2, Appli
   20        35      43.8     31  2  US-08-44[illegible]-873-2   Sequence 2, Appli
   21        35      43.8     31  4  US-0[illegible]-452B-2      Sequence 2, Appli
   22        35      43.8    217  2  US-08-62[illegible]-120A-25 Sequence 25, Appli
   23        35      43.8    314  2  US-08-52[illegible]-12-6    Sequence 6, Appli
   24        35      43.8    421  4  US-08-7[illegible]-23-5     Sequence 5, Appli
   25        35      43.8    513  2  US-08-67[illegible]-166A-4  Sequence 4, Appli
   26        35      43.8    513  2  US-08-67[illegible]-166A-5  Sequence 5, Appli
   27        35      43.8    584  1  US-08-2[illegible]-22A-8    Sequence 8, Appli
```

```
   28        35      43.8    584  5  PCT-US9[illegible]-744A-8   Sequence 8, Appli
   29        35      43.8    697  2  US-08-4[illegible]-396A-[illegible] Sequence 50, Appli
   30        35      43.8    697  3  US-08-[illegible]           Sequence 5[illegible], Appl
   31        35      43.8    697  4  US-08-46[illegible]-371A-[illegible] Sequence 50, Appl
   32        35      43.8    697  4  US-08-46[illegible]         Sequence 50, Appl
   33        35      43.8    711  1  US-08-[illegible]           Sequence 8, Appli
   34        35      43.8    711  1  US-08-[illegible]           Sequence 2, Appli
   35        35      43.8    711  2  US-08-[illegible]           Sequence 1, Appli
   36        35      43.8    711  2  US-08-[illegible]           Sequence 2, Appli
   37        35      43.8    711  4  US-09-[illegible]           Sequence 2, Appli
   38        35      43.8    711  5  PCT-US9[illegible]          Sequence 2, Appli
   39        35      43.8    723  1  US-07-[illegible]           Sequence 1, Appli
   40        35      43.8    723  1  US-08-[illegible]           Sequence 1, Appli
   41        35      43.8    723  1  US-08-[illegible]-947B-2    Sequence 2, Appli
   42        35      43.8    723  1  US-08-[illegible]-947B-3    Sequence 3, Appli
   43        35      43.8    723  1  US-08-404-64[illegible]     Sequence 1, Appli
   44        35      43.8    723  4  US-08-[illegible]           Sequence 3, Appli
   45        35      43.8    728  1  US-07-[illegible]-333A-2    Sequence 2, Appli


                                   ALIGNMENTS


RESULT   1
US-08-538-960-3
; Sequence 3, Application US/08538960
; Patent No. 5872[illegible]
;  GENERAL INFORMATION:
;    APPLICANT:  Stocco, Douglas M.
;    APPLICANT:  Clark, Barbara J.
;    TITLE OF INVENTION: [illegible]
;    TITLE OF INVENTION: [illegible]ENESIS
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE: [illegible]ump, Strauss, [illegible] Feld [illegible]
;      STREET: [illegible] Avenue
;      CITY:  Austin
;      STATE:  TX
;      COUNTRY:  U.S.A.
;      ZIP:  787[illegible]
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.[illegible] Version #1.[illegible]
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/5[illegible]
;      FILING DATE:  Concurrently Herewith
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Mayfield, Denise L.
;      REGISTRATION NUMBER:  33,732
;      REFERENCE/DOCKET NUMBER:  13[illegible]2/DLM
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  512/499-6200
;      TELEFAX:  512/499-6290
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  14 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-538-960-3


  Query Match            100.0%;  Score 80;  DB 2;  Length 14;
  Best Local Similarity  100.0%;  Pred. No. 2.8e-07;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AEHGPTCMVLHPLA 14
              ||||||||||||||
Db          1 AEHGPTCMVLHPLA 14
```

```
RESULT 2
US-08-659-254-3
; Sequence 3, Application US/08659254
; Patent No. 61945[illegible]
; GENERAL INFORMATION:
;   APPLICANT: [illegible] M
;   APPLICANT: Clark, [illegible] Barbara
;   TITLE OF INVENTION: COMPOSITIONS AND METH[illegible]
;   TITLE OF INVENTION: [illegible]
;   NUMBER OF SEQUENCES: [illegible]
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: [illegible]
;     STREET: [illegible]
;     CITY: Austin
;     STATE: TX
;     COUNTRY: [illegible]
;     ZIP: 78[illegible]
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: [illegible]/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #[illegible]
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/659,254
;     FILING DATE: [illegible]
;     CLASSIFICATION: [illegible]
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: US 08/[illegible]
;     FILING DATE: [illegible]
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Mayfield, Denise L.
;     REGISTRATION NUMBER: [illegible]
;     REFERENCE/DOCKET NUMBER: [illegible]
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: 512/[illegible]
;     TELEFAX: 512/[illegible]
; INFORMATION FOR SEQ ID NO: [illegible]
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 11 amino acids
;     TYPE: amino acid
;     STRANDEDNESS: [illegible]
;     TOPOLOGY: linear
;   MOLECULE TYPE: peptide
US-08-659-254-3

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches [illegible]; Conservative [illegible]; Mismatches [illegible]

Qy          1 ALHGPTCMVLH[illegible]A 1[illegible]
              [illegible]
Db          1 AFHGPTCMVLH[illegible]A 1[illegible]


RESULT 3
US-08-659-254-2
; Sequence 2, Application US/08659254
; Patent No. [illegible]94[illegible]
; GENERAL INFORMATION
;   APPLICANT: [illegible] M
;   APPLICANT: Clark, Dr. Barbara
;   TITLE OF INVENTION: [illegible] AND METH[illegible]
;   TITLE OF INVENTION: [illegible]
;   NUMBER OF SEQUENCES: [illegible]
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: [illegible]
;     STREET: [illegible]
;     CITY: Austin
;     STATE: TX
;     COUNTRY: U.S.A.
```

```
;     [illegible]
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: [illegible]
;     COMPUTER: [illegible]
;     OPERATING SYSTEM: [illegible]
;     SOFTWARE: [illegible] Version #[illegible]
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: [illegible]
;     FILING DATE: [illegible]
;     CLASSIFICATION: [illegible]
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: [illegible]
;     FILING DATE: [illegible]
;   ATTORNEY/AGENT INFORMATION:
;     NAME: [illegible]
;     REGISTRATION NUMBER: [illegible]
;     REFERENCE/DOCKET NUMBER: [illegible]
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: [illegible]
;     TELEFAX: [illegible]
; INFORMATION FOR SEQ ID NO: [illegible]
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: [illegible] amino acids
;     TYPE: amino acid
;     STRANDEDNESS: [illegible]
;     TOPOLOGY: [illegible]
;   MOLECULE TYPE: peptide
US-[illegible]

  Query Match [illegible] Score [illegible] DB 4; Length [illegible]
  Best Local Similarity [illegible] Pred. No. [illegible]
  Matches [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible]

Qy          [illegible] A[illegible]MV[illegible]A 1[illegible]

Db          [illegible] A[illegible]MV[illegible]


RESULT 4
[illegible]
; Sequence [illegible], Application [illegible]
; Patent No. [illegible]
; GENERAL INFORMATION:
;   APPLICANT: [illegible] M
;   APPLICANT: [illegible]
;   TITLE OF INVENTION: [illegible] AND ME[illegible]
;   TITLE OF INVENTION: [illegible]
;   NUMBER OF SEQUENCES: 1[illegible]
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: [illegible]
;     STREET: [illegible] Austin
;     CITY: Austin
;     STATE: TX
;     COUNTRY: [illegible]
;     ZIP: [illegible]
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM [illegible]
;     OPERATING SYSTEM: [illegible]
;     SOFTWARE: [illegible] Version #[illegible]
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: [illegible]
;     FILING DATE: [illegible]
;     CLASSIFICATION: [illegible]
;   ATTORNEY/AGENT INFORMATION:
;     NAME: [illegible]
;     REGISTRATION NUMBER: [illegible]
;     REFERENCE/DOCKET NUMBER: [illegible]
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: [illegible]
;     TELEFAX: [illegible]
; INFORMATION FOR SEQ ID NO: [illegible]
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  284 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-538-960-2


  Query Match              100.0%;  Score 80;  DB 2;  Length 284;
  Best Local Similarity    100.0%;  Pred. No. 6.7e-06;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AEHGPTCMVLHPLA 14
              ||||||||||||||
Db        217 AEHGPTCMVLHPLA 230


RESULT   5
US-08-410-540-2
; Sequence 2, Application US/08410540
; Patent No. 5807678
;  GENERAL INFORMATION:
;     APPLICANT:  Miller, Walter L.
;     APPLICANT:  Lin, Dong
;     APPLICANT:  Strauss III, Jerome F
;     TITLE OF INVENTION:  IDENTIFICATION OF GENE MUTATIONS
;     TITLE OF INVENTION:  ASSOCIATED WITH CONGENITAL LIPOID ADRENAL HYPERPLASIA
;     NUMBER OF SEQUENCES:  30
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Cooley Godward Castro Huddleson & Tatum
;       STREET:  5 Palo Alto Square
;       CITY:  Palo Alto
;       STATE:  CA
;       COUNTRY:  US
;       ZIP:  94306-2155
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/410,540
;       FILING DATE:  23-MAR-1995
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Neeley, Richard L.
;       REGISTRATION NUMBER:  30,092
;       REFERENCE/DOCKET NUMBER:  UCAL 234/00US
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  415 853 5070
;       TELEFAX:  415 857 0663
;       TELEX:  380816COOLEYPA
;   INFORMATION FOR SEQ ID NO:  2:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  285 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-410-540-2


  Query Match              100.0%;  Score 80;  DB 1;  Length 285;
  Best Local Similarity    100.0%;  Pred. No. 6.7e-06;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AEHGPTCMVLHPLA 14
              ||||||||||||||
Db        218 AEHGPTCMVLHPLA 231


RESULT   6
US-08-659-254-18
; Sequence 18, Application US/08659254
; Patent No. 6194556
;  GENERAL INFORMATION:
;     APPLICANT:  Stocco, Douglas M.
;     APPLICANT:  Clark, Dr. Barbara J.
;     TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR REGULATION
;     TITLE OF INVENTION:  OF STEROIDOGENESIS
;     NUMBER OF SEQUENCES:  19
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Akin, Gump, Strauss, Hauer & Feld, L.L.P.
;       STREET:  1900 Frost Bank Plaza, 816 Congress Avenue
;       CITY:  Austin
;       STATE:  TX
;       COUNTRY:  U.S.A.
;       ZIP:  78701
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/659,254
;       FILING DATE:  07 JUN 1996
;       CLASSIFICATION:  530
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/538,960
;       FILING DATE:  04-OCT-1995
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Mayfield, Denise L.
;       REGISTRATION NUMBER:  33,732
;       REFERENCE/DOCKET NUMBER:  43375.0006
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  512/499 6200
;       TELEFAX:  512/499-6290
;   INFORMATION FOR SEQ ID NO:  18:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  285 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
US-08-659-254-18


  Query Match              100.0%;  Score 80;  DB 4;  Length 285;
  Best Local Similarity    100.0%;  Pred. No. 6.7e-06;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AEHGPTCMVLHPLA 14
              ||||||||||||||
Db        218 AEHGPTCMVLHPLA 231


RESULT   7
US-08-538-960-7
; Sequence 7, Application US/08538960
; Patent No. 5872240
;  GENERAL INFORMATION:
;     APPLICANT:  Stocco, Douglas M.
;     APPLICANT:  Clark, Barbara J.
;     TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR
;     TITLE OF INVENTION:  REGULATION OF STEROIDOGENESIS
;     NUMBER OF SEQUENCES:  14
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Akin, Gump, Strauss, Hauer & Feld L.L.P.
;       STREET:  1900 Frost Bank Plaza 816 Congress Avenue
;       CITY:  Austin
;       STATE:  TX
;       COUNTRY:  U.S.A.
;       ZIP:  78701
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER: [illegible]
;      FILING DATE: [illegible]
;      CLASSIFICATION: [illegible]
;    ATTORNEY/AGENT INFORMATION:
;      NAME: Ma[illegible]
;      REGISTRATION NUMBER: [illegible]
;      REFERENCE/DOCKET NUMBER: [illegible]
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE: [illegible]
;      TELEFAX: [illegible]
;  INFORMATION FOR SEQ ID NO: [illegible]
;    SEQUENCE CHARACTERISTICS:
;      LENGTH: [illegible] amino acids
;      TYPE: amino acid
;      STRANDEDNESS: [illegible]
;      TOPOLOGY: linear
;    MOLECULE TYPE: peptide
[illegible]8-[illegible]7

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches [illegible]; Conservative [illegible]; Mismatches [illegible]

Qy          1 AFHGPTCMV [illegible]
              |||||||||
Db          1 AFHGPTCMV [illegible]


RESULT 8
US-08-659-254-7
; Sequence 7, Application US/0865[illegible]
; Patent No. 6194[illegible]
; GENERAL INFORMATION:
;    APPLICANT: [illegible] M
;    APPLICANT: [illegible]
;    TITLE OF INVENTION: COMPOSITION AND ME[illegible]
;    TITLE OF INVENTION: [illegible]
;    NUMBER OF SEQUENCES: [illegible]
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE: [illegible]
;      STREET: [illegible]
;      CITY: Austin
;      STATE: TX
;      COUNTRY: [illegible]
;      ZIP: 787[illegible]
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE: [illegible]
;      COMPUTER: [illegible]
;      OPERATING SYSTEM: [illegible]
;      SOFTWARE: PatentIn Release #[illegible], Version #[illegible]
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER: [illegible]
;      FILING DATE: [illegible]
;      CLASSIFICATION: [illegible]
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER: [illegible]
;      FILING DATE: [illegible]
;    ATTORNEY/AGENT INFORMATION:
;      NAME: Mayfield, De[illegible]
;      REGISTRATION NUMBER: [illegible]
;      REFERENCE/DOCKET NUMBER: [illegible]
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE: [illegible]
;      TELEFAX: [illegible]
;  INFORMATION FOR SEQ ID NO: [illegible]
;    SEQUENCE CHARACTERISTICS:
;      LENGTH: 9 amino acids
;      TYPE: amino acid
;      STRANDEDNESS: [illegible]
;      TOPOLOGY: linear
;    MOLECULE TYPE: peptide
```

```
[illegible]

  Query Match [illegible] Length [illegible]
  Best Local Similarity [illegible]
  Matches [illegible] Mismatches [illegible] Indels [illegible] Gaps [illegible]

Qy          1 [illegible]CMV [illegible]

Db          1 [illegible]CMV [illegible]


RESULT [illegible]
[illegible]
; Sequence [illegible], Application [illegible]
; Patent No. [illegible]
; GENERAL INFORMATION:
;    APPLICANT: [illegible] MATTHEW [illegible]
;    APPLICANT: [illegible]
;    APPLICANT: [illegible]
;    TITLE OF INVENTION: [illegible] Genes and [illegible]
;    NUMBER OF SEQUENCES: [illegible]
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE: [illegible]
;      STREET: Four Embarcadero Center, Suite [illegible]
;      CITY: [illegible]
;      STATE: CA
;      COUNTRY: [illegible]
;      ZIP: [illegible]
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE: Floppy disk
;      COMPUTER: [illegible] compatible
;      OPERATING SYSTEM: [illegible]
;      SOFTWARE: PatentIn Release #[illegible], Version #[illegible]
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER: [illegible]
;      FILING DATE: [illegible]
;      CLASSIFICATION: [illegible]
;    ATTORNEY/AGENT INFORMATION:
;      NAME: [illegible]
;      REGISTRATION NUMBER: [illegible]
;      REFERENCE/DOCKET NUMBER: [illegible]
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE: [illegible]
;      TELEFAX: [illegible]
;  INFORMATION FOR SEQ ID NO: [illegible]
;    SEQUENCE CHARACTERISTICS:
;      LENGTH: [illegible]
;      TYPE: [illegible]
;      STRANDEDNESS: [illegible]
;      TOPOLOGY: [illegible]
;    MOLECULE TYPE: [illegible]
[illegible]

  Query Match [illegible] Length [illegible]
  Best Local Similarity [illegible] Pred. No. [illegible]
  Matches [illegible] Mismatches [illegible] Indels [illegible] Gaps [illegible]

Qy          [illegible]

Db          [illegible]NAIVE[illegible]


RESULT [illegible]
[illegible]
; Sequence [illegible], Application [illegible]
; Patent No. [illegible]
; GENERAL INFORMATION:
;    APPLICANT: [illegible] MATTHEW [illegible]
;    APPLICANT: [illegible] V
;    APPLICANT: [illegible]
;    TITLE OF INVENTION: [illegible] Genes and [illegible]
;     NUMBER OF SEQUENCES:  19
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Flehr, Hohbach, Test, Albritton & Herbert
;       STREET:  Four Embarcadero Center, Suite 3400
;       CITY:  San Francisco
;       STATE:  CA
;       COUNTRY:  US
;       ZIP:  94111
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM.  PC DOS/MS DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/656,055
;       FILING DATE:
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/540,406
;       FILING DATE:
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Rowland, Bertram I
        REGISTRATION NUMBER:  20015
        REFERENCE/DOCKET NUMBER:  a60190-1
      TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  415-781-1989
;       TELEFAX:  415-398-3249
;   INFORMATION FOR SEQ ID NO:  6:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  1285 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-656-055-6

  Query Match             47.5%;  Score 38;  DB 3;  Length 1285;
  Best Local Similarity   58.3%;  Pred. No. 2.4e-02;
  Matches    7;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       3 HGPTCMVLHPLA 14
             | |   |||| |
Db     139 HDPNASVLHPQA 150


RESULT 11
US-08-954-668-6
; Sequence 6, Application US/08954668
; Patent No. 6172200
  GENERAL INFORMATION:
      APPLICANT:  SCOTT, MATHEW P
      APPLICANT:  GOODRICH, LISA V
;     APPLICANT:  JOHNSON, RONALD L
;     TITLE OF INVENTION:  Patched Genes and their Use
;     NUMBER OF SEQUENCES:  19
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Foley, Hoag & Eliot
;       STREET:  One Post Office Square
;       CITY:  Boston
;       STATE:  MA
;       COUNTRY:  USA
;       ZIP:  02109
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  ASCII(text)
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/954,668
;       FILING DATE:  20-Oct-1997
;       CLASSIFICATION:  530
;     ATTORNEY/AGENT INFORMATION:
```

```
;       NAME.  Vincent, Matthew P.
;       REGISTRATION NUMBER:  36709
;       REFERENCE/DOCKET NUMBER:  SUV-00[illegible]
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  617 832-1000
;       TELEFAX:  617-832 7000
;   INFORMATION FOR SEQ ID NO:  6:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  1285 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-954-668-6

  Query Match             47.5%;  Score 38;  DB 4;  Length 1285;
  Best Local Similarity   58.3%;  Pred. No. 2.4e-02;
  Matches    7;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       3 HGPTCMVLHPLA 14
             | |   ||||
Db     139 HDPNASVLHPQA 150


RESULT 12
PCT-US95-13233-6
; Sequence 6, Application PC/TUS9513233
;   GENERAL INFORMATION:
;     APPLICANT.  THE BOARD OF TRUSTEES OF THE LELAND STANFORD JUNIOR UNIVERSITY
;     TITLE OF INVENTION.  Patched Genes and their Use
;     NUMBER OF SEQUENCES:  19
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Flehr, Hohbach, Test, Albritton & Herbert
;       STREET:  Four Embarcadero Center, Suite 3400
;       CITY:  San Francisco
;       STATE:  CA
;       COUNTRY:  US
;       ZIP:  94111
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM.  PC DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/US95/13233
;       FILING DATE:  06-OCT-1990
;       CLASSIFICATION:
      ATTORNEY/AGENT INFORMATION:
,       NAME:  Rowland, Bertram I
;       REGISTRATION NUMBER:  20015
;       REFERENCE/DOCKET NUMBER:  a60190
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  415-781-1989
,       TELEFAX.  415 398 3249
;   INFORMATION FOR SEQ ID NO:  6:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  1285 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
,     MOLECULE TYPE:  protein
PCT-US95-13233-6

  Query Match             47.5%,  Score 38;  DB 5;  Length 1285;
  Best Local Similarity   58.3%;  Pred. No. 2.4e-02;
  Matches    7;  Conservative    0,  Mismatches    5;  Indels    0;  Gaps    0;

Qy       3 HGPTCMVLHPLA 14
             | |   |||| |
Db     139 HDPNASVLHPQA 150
```

```
RESULT 13
[illegible]-268-[illegible]-3
; Sequence [illegible], Application [illegible]
; Patent No. 62[illegible]
; GENERAL INFORMATION:
;   APPLICANT: [illegible] M
;   APPLICANT: [illegible] A
;   TITLE OF INVENTION: [illegible]
;   FILE REFERENCE: [illegible]
;   CURRENT APPLICATION NUMBER: [illegible]
;   CURRENT FILING DATE: [illegible]
;   PRIOR APPLICATION NUMBER: [illegible]
;   PRIOR FILING DATE: [illegible]
;   NUMBER OF SEQ ID NOS: [illegible]
;   SOFTWARE: PatentIn Ver. [illegible]
; SEQ ID NO [illegible]
;  LENGTH: 1286
;  TYPE: [illegible]
;  ORGANISM: Drosophila [illegible]
[illegible]-268-[illegible]-3


  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches [illegible] Conservative [illegible] Mismatches [illegible]

Qy        [illegible]CMV[illegible] 14
              [illegible]
Db        [illegible]NASV[illegible]


RESULT 14
[illegible]-4[illegible]-6
; Sequence [illegible], Application [illegible]
; Patent No. [illegible]16[illegible]
; GENERAL INFORMATION:
;     APPLICANT: [illegible] W
;     APPLICANT: McM[illegible], Andrew [illegible]
;     APPLICANT: [illegible]
;     APPLICANT: [illegible] A
;     APPLICANT: [illegible]
;     TITLE OF INVENTION: [illegible]
;     TITLE OF INVENTION: [illegible]
;     NUMBER OF SEQUENCES: [illegible]
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE: [illegible]
;       STREET: One [illegible] Square
;       CITY: Boston
;       STATE: MA
;       COUNTRY: USA
;       ZIP: [illegible]
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE: [illegible] disk
;       COMPUTER: IBM [illegible] compatible
;       OPERATING SYSTEM: [illegible]/MS-[illegible]
;       SOFTWARE: PatentIn Release #[illegible] Version #1
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER: [illegible]
;       FILING DATE: [illegible]
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER: [illegible]
;       FILING DATE: [illegible] MAY [illegible]
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER: [illegible]
;       FILING DATE: [illegible] DEC [illegible]
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER: [illegible]
;       FILING DATE: [illegible] DEC [illegible]
;     ATTORNEY/AGENT INFORMATION:
;       NAME: Vincent [illegible]
;       REGISTRATION NUMBER: [illegible]
;       REFERENCE/DOCKET NUMBER: HMV[illegible]
```

```
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE: [illegible]
;       TELEFAX: [illegible]
;   INFORMATION FOR SEQ ID NO: [illegible]
;     SEQUENCE CHARACTERISTICS:
;       LENGTH: [illegible]
;       TYPE: [illegible]
;       [illegible]
;     MOLECULE TYPE: [illegible]
[illegible]


  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches [illegible] Mismatches [illegible] Indels [illegible] Gaps [illegible]

Qy        [illegible]
              [illegible]
Db        [illegible]NA[illegible]


RESULT 15
[illegible]
; Sequence [illegible], Application [illegible]A
; Patent No. [illegible]
; GENERAL INFORMATION:
;     APPLICANT: [illegible]
;     APPLICANT: Yama[illegible] M[illegible]
;     APPLICANT: [illegible]
;     APPLICANT: Yama[illegible] K[illegible]
;     TITLE OF INVENTION: [illegible]
;     NUMBER OF SEQUENCES: [illegible]
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE: [illegible]
;       STREET: [illegible]
;       CITY: [illegible]
;       STATE: [illegible]
;       COUNTRY: [illegible]
;       ZIP: [illegible]
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE: [illegible] DISK
;       COMPUTER: IBM [illegible]
;       OPERATING SYSTEM: [illegible]/MS-[illegible]
;       SOFTWARE: [illegible] #[illegible]
;       SOFTWARE: Version #[illegible]
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER: [illegible]A
;       FILING DATE: [illegible] MAY [illegible]
;     ATTORNEY/AGENT INFORMATION:
;       NAME: W. PATRICK BENGTSSON
;       REGISTRATION NUMBER: [illegible]
;       REFERENCE/DOCKET NUMBER: [illegible]
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE: [illegible]
;       TELEFAX: [illegible]
;       TELEX: [illegible]
;   INFORMATION FOR SEQ ID NO: [illegible]
;     SEQUENCE CHARACTERISTICS:
;       LENGTH: [illegible]
;       TYPE: [illegible]
;       STRANDEDNESS: [illegible]
;       [illegible]
;     MOLECULE TYPE: [illegible]
;     ORIGINAL SOURCE:
;       ORGANISM: [illegible]
;       [illegible]
;       [illegible]
;     FEATURE:
;       LOCATION: [illegible]
;       NAME/KEY: [illegible]
;     FEATURE:
;       LOCATION: [illegible]
;       NAME/KEY: [illegible]
;     FEATURE:
;        LOCATION:  21
;        NAME/KEY:  Xaa:Ser or Glu
;     FEATURE:
;        LOCATION:  58
;        NAME/KEY:  Xaa:Arg or Met
;     FEATURE:
;        LOCATION:  59
;        NAME/KEY:  Xaa:Val or Ala
;     FEATURE:
;        LOCATION:  61
;        NAME/KEY:  Xaa:Gln or Glu
;     FEATURE:
;        LOCATION:  69
;        NAME/KEY:  Xaa:Ala or Ser
;     FEATURE:
;        LOCATION:  77
;        NAME/KEY:  Xaa:Tyr or Asp
US-08-451-409A-1

  Query Match             46.2%;  Score 37;  DB 1;  Length 21[illegible];
   est Local Similarity   50.0%;  Pred. No. [illegible];
    tches    7;  Conservative    3;  Mismatches    4;  Indels    [illegible];  Gaps    [illegible];

Qy       1 AEHGPTCMVLHPLA 14
           |||||   :|: :|
Db     125 AEHGPLSILLNAVA 138


Search completed: August 22, 2001, 14:[illegible]
Job time: 364 sec
```

```
                    GenCore version 4.5
          Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on:          August 22, 2001, 14:25:22   Search time 64.49 Seconds
                                              (without alignments)
                                              16.537 Million cell updates/sec

Title:           US-08-659-254-3
Perfect score:   80
Sequence:        1 AEHGPTCMVLHPLA 14

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters:      219241

[illegible]imum DB seq length: 0
[illegible]mum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       PIR_68:*
                1:  pir1:*
                2:  pir2:*
                3:  pir3:*
                4:  pir4:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                              SUMMARIES
                   %
Result           Query
   No.   Score   Match Length DB  ID                                    Description
-------------------------------------------------------------------------------------
     1      80   100.0    284  2  A55455                            steroidogenic acut
     2      80   100.0    285  2  I38248                            steroidogenic acut
     3      80   100.0    362  2  JC5386                            steroidogenic acut
     4      72    90.0    285  2  JC4315                            steroidogenic acut
     5      45    56.2    730  2  A75486                            hypothetical prote
     6    44.5    55.6    269  2  B84013                            hypothetical prote
     7    43.5    54.4    325  1  S15752                            acid phosphatase (
     8    43.5    54.4    327  1  JN0787                            acid phosphatase (
     9    43.5    54.4    327  1  A41720                            acid phosphatase (
    10      43    53.8    670  2  C49934                            transketolase (E
    11      42    52.5    549  2  JN0552                            triacylglycerol li
    12      41    51.2    246  2  A60049                            growth associated
    13      41    51.2    317  2  E83380                            probable transcrip
    14      41    51.2    442  2  E82960                            probable oxidoredu
    15      41    51.2    515  2  E86610                            axial filament pro
    16      41    51.2    515  2  D72015                            [illegible]
    17      41    51.2    638  2  D86427                            protein F15O4.27 [
    18      41    51.2   1849  2  T00415                            hypothetical prote
    19    40.5    50.6    340  1  A46096                            acid phosphatase (
    20      40    50.0    396  2  F86183                            hypothetical prote
    21      40    50.0    431  2  S49821                            PRL protein - Aca
    22      40    50.0    453  2  T01114                            hypothetical prote
    23      40    50.0    548  2  S32615                            triacylglycerol li
    24      40    50.0    684  2  E75434                            probable [illegible]
    25      40    50.0   4543  1  A53142                            alpha 2 macroglobu
    26      39    48.8    228  2  T26542                            hypothetical prote
    27      39    48.8    372  2  A82264                            UDP N-acetylglucos
    28      39    48.8    459  2  H64667                            conserved hypothet
    29      39    48.8    459  2  A71819                            hypothetical prote
    30      39    48.8    512  2  D71468                            probable axial fil
    31      39    48.8    512  2  F81749                            ribonuclease G (50
    32      39    48.8   1309  2  F96509                            protein F27F5.1 [
    33      39    48.8   1872  2  I30888                            vitellogenin - Ath
    34      38    47.5    193  2  B71878                            recombination prot
    35      38    47.5    193  2  E64[illegible]                            recombinational DN
    36      38    47.5    281  2  T04760                            hypothetical prote
    37      38    47.5    281  2  D85223                            probable LTR retro
    38      38    47.5    662  2  T28364                            hypothetical prote
    39      38    47.5    718  2  H75485                            hypothetical prote
    40      38    47.5   1251  2  B86194                            hypothetical prote
    41      38    47.5   1299  1  S06113                            membrane protein p
    42      38    47.5   1433  2  A71444                            probable LTR retro
    43      37    46.2    185  2  C82711                            hypothetical prote
    44      37    46.2    274  2  D82784                            phosphomethylpyrim
    45      37    46.2    275  2  S27612                            nifU protein - Pse

                              ALIGNMENTS

RESULT    1
A55455
steroidogenic acute regulatory protein precursor, mitochondrial - mouse
C;Species: Mus musculus (house mouse)
C;Date: 10-Feb-1995 #sequence_revision 10-Feb-1995 #text_change 08-Oct-1999
C;Accession: A55455
R;Clark, B.J.; Wells, J.; King, S.R.; Stocco, D.M.
J. Biol. Chem. 269, 28314-28322, 1994
A;Title: The purification, cloning, and expression of a novel luteinizing hormone-ind
ry protein (StAR)
A;Reference number: A55455; MUID:9505061[illegible]
A;Accession: A55455
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-284 <CLA>
A;Cross-references: GB:L36062; GB:[illegible]; NID:g1200242; PIDN:AAB94783.1; P
C;Keywords: mitochondrion; phosphoprotein
F;1-47/Domain: transit peptide (mitochondrion) #status predicted <TPP>

  Query Match             100.0%;  Score 80;  DB 2;  Length 284;
  Best Local Similarity   100.0%;  Pred. No. 4.1e-06;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 AEHGPTCMVLHPLA 14
              ||||||||||||||
Db      217 AEHGPTCMVLHPLA 230


RESULT    2
I38248
steroidogenic acute regulatory protein - human
C;Species: Homo sapiens (man)
C;Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 05-Nov-1999
C;Accession: I38248; I38896
R;Sugawara, T.; Lin, D.; Holt, J.A.; Martin, K.O.; Javitt, N.B.; Miller, W.L.; Straus
Biochemistry 34, 12506-12512, 1995
A;Title: Structure of the human steroidogenic acute regulatory protein (StAR) gene: S
A;Reference number: I38248; MUID:96038208
A;Accession: I38248
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-285 <RES>
A;Cross-references: EMBL:U29105; NID:[illegible]; PIDN:[illegible]; PID:g1041698
R;Sugawara, T.; Holt, J.A.; Driscoll, D.; Strauss III, J.F.; Lin, D.; Miller, W.L.; P
Proc. Natl. Acad. Sci. U.S.A. 92, 4778-4782, 1995
A;Title: Human steroidogenic acute regulatory protein: functional activity in COS-1 c
A;Reference number: I38896; MUID:95281[illegible]
A;Accession: I38896
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-285 <FL2>
```

```
A;Cross-references: EMBL [illegible]; NID [illegible]; PIDN:AA[illegible]
C;Genetics:
A;Gene: STAR
A;Cross-references: [illegible] STAR; GDB [illegible]; MIM [illegible]
A;Map position: 8p11 [illegible] 8p1[illegible]
A;Introns: [illegible]

  Query Match [illegible]
  Best Local Similarity [illegible]; Pred. N[illegible]
  Matches [illegible]4; Conservative [illegible]; Mismatches [illegible]

Qy          1 AFHGPTCMV[illegible]LA [illegible]
              ||||||||||  ||
Db        218 AFHGPTCMV[illegible]LA [illegible]


RESULT      3
[illegible]86
steroidogenic acute regulatory protein - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 04-Jun-1997 #sequence_revision 18-Jul-1997 #text_change [illegible]
C;Accession: JC5386
R;Lee, H.; Ahn, R.S.; Kwon, H.; Soh, [illegible]
Biochem. Biophys. Res. Commun. [illegible]
A;Title: Nucleotide sequence of rat steroidogenic [illegible]
A;Reference number: JC5386; MUID:[illegible]
A;Accession: JC5386
A;Molecule type: mRNA
A;Residues: 1-36[illegible]
A;Cross-references: [illegible]
C;Comment: This protein [illegible]

  Query Match [illegible]
  Best Local Similarity [illegible]; Pred. N[illegible]
  Matches   11; Conservative [illegible]; Mismatches [illegible]

Qy          1 AFHGPTCMVLHPLA [illegible]
              [illegible]
Db        295 AFHGPTCMVLHPLA [illegible]


RESULT      4
JC4315
[illegible]idogenic acute regulatory protein - [illegible]
[illegible]ecies: Bos primigenius taurus (cattle)
C;Date: 29-Nov-199[illegible] #sequence_revision [illegible]
C;Accession: JC4315
R;Hartung, S.; Rust, W.; [illegible]; M[illegible]
Biochem. Biophys. Res. Commun. [illegible]
A;Title: Molecular cloning and [illegible] expression [illegible]
A;Reference number: JC4315; MUID:[illegible]
A;Accession: JC4315
A;Molecule type: mRNA
A;Residues: 1-285 [illegible]
C;Comment: This protein [illegible]
C;Genetics:
A;Gene: STAR
F;226-264/Region: [illegible]

  Query Match [illegible]
  Best Local Similarity [illegible]; Pred. N[illegible]
  Matches [illegible]; Conservative [illegible]; Mismatches [illegible]

Qy          1 AFHGPTCMV[illegible]LA [illegible]
              ||||||||||      |
Db        218 AFHGPTCMV[illegible]LA [illegible]


RESULT      5
A[illegible]

hypothetical protein [illegible] Deinococcus radiodurans [illegible]
C;Species: Deinococcus radiodurans
C;Date: [illegible] #sequence_revision [illegible] #text_change [illegible]
C;Accession: A[illegible]
R;White, O.; Eisen, J.A.; Heidelberg, J.F.; [illegible]
[illegible]
[illegible] Smith, H.O.; Venter, J.C.; Fraser, C.M.
[illegible]
A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1
A;Reference number: A[illegible]; MUID:[illegible]
A;Accession: A[illegible]
A;Status: preliminary
A;Molecule type: DNA
A;Residues: [illegible]
A;Cross-references: GB:AE[illegible]; GB:AE[illegible]; NID:[illegible]; PIDN:AAF[illegible]; PID:[illegible]
A;Experimental source: strain R1
C;Genetics:
A;Gene: [illegible]
A;Map position: [illegible]
C;Superfamily: Deinococcus radiodurans hypothetical protein DR[illegible]

  Query Match [illegible]
  Best Local Similarity [illegible]; Pred. No. [illegible]
  Matches [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible]

Qy [illegible]
[illegible]
Db [illegible]


RESULT [illegible]
[illegible]
hypothetical protein [illegible] (imported) - [illegible] halodurans (strain [illegible])
C;Species: [illegible]
C;Date: [illegible] #sequence_revision [illegible] #text_change [illegible]
C;Accession: [illegible]
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, [illegible]; Sasaki, [illegible]; Masui, N.; Fuji, F.; [illegible]
Nucleic Acids Res. [illegible]
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans [illegible]
A;Reference number: A[illegible]; MUID:[illegible]
A;Accession: [illegible]
A;Status: preliminary
A;Molecule type: DNA
A;Residues: [illegible]
A;Cross-references: [illegible] BA[illegible]; NID:[illegible]; PIDN:BAB[illegible]; GSPDB:[illegible]
A;Experimental source: [illegible]
C;Genetics:
A;Gene: BH[illegible]

  Query Match [illegible]
  Best Local Similarity [illegible]; Pred. No. [illegible]
  Matches [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible]

Qy [illegible] MVF[illegible]
[illegible]
Db [illegible]


RESULT [illegible]
[illegible]
acid phosphatase (EC [illegible]) ACP5 precursor - human
N;Alternate names: [illegible] acid phosphatase (TRAP); tartrate-resistant acid phosphatase
C;Species: Homo sapiens (man)
C;Date: [illegible] #sequence_revision [illegible] #text_change [illegible] Mar [illegible]
C;Accession: [illegible]; A[illegible]; [illegible]; [illegible]; A4[illegible]; A34638; [illegible]; A33781; B4[illegible];
R;[illegible]; Ketcham, C.M.; [illegible]; Roman, J.A.; Gross, A.V
[illegible] Biochem. [illegible]
A;Title: [illegible] phosphatase: [illegible] expression and [illegible]
A;Reference number: [illegible]; MUID:[illegible]
A;Accession: [illegible]
A;Molecule type: mRNA
```

```
A;Residues: 1 325 <LOR>
A;Cross-references: EMBL:X14618; NID:g347[illegible] [illegible]
A;Experimental source: placenta
R;Ketcham, C.M.; Roberts, R.M.; Simmen, R.C.M  Nick, H.S.
J. Biol. Chem. 264, 557-563, 1989
A;Title: Molecular cloning of the type 5, iron-containing, tartrate-resistant acid phosp
A;Reference number: A31812; MUID:89079710
A;Accession: A31812
A;Molecule type: mRNA
A;Residues: 1-44,'GP',47-176,'LI',181-325 <KET>
A;Cross-references: GB:J04430; GB:M1953[illegible]; NID:g17[illegible]; PIDN:AAA[illegible]; PID:g[illegible]
A;Experimental source: placenta
R;Cassady, A.I.; King, A.G.; Cross, N.C.P.; Hume, D.A.
Gene 130, 201-207, 1993
A;Title: Isolation and characterization of the genes encoding mouse and human type 5 aci
A;Reference number: JN0787; MUID:93366175
A;Accession: PN0621
A;Molecule type: DNA
A;Residues: 1-245 <CAS>
A;Cross-references: GB:X67123
R;Hayman, A.R.; Warburton, M.J.; Pringle, J.A.S.; Coles, B.; Chambers, T.
Biochem. J. 261, 601-609, 1989
[illegible]tle: Purification and characterization of a tartrate-resistant acid phosphatase from
[illegible]ference number: S05251; MUID:89374153
A;Accession: S05251
A;Molecule type: protein
A;Residues: 22-46,'Q',48-49,'W',51-64;18[illegible] 'A' <HAY>
A;Experimental source: osteclastoma
R;Janckila, A.J.; Latham, M.D.; Lam, K.W.; Chow, K.C.; Li, C.Y.; Yam, L.T.
Clin. Biochem. 25, 437-443, 1992
A;Title: Heterogeneity of hairy cell tartrate resistant acid phosphatase.
A;Reference number: A49133; MUID:93121546
A;Accession: A49133
A;Molecule type: protein
A;Residues: 22-46,'G',48-49,'X',51-55;18[illegible] <JA4>
A;Experimental source: spleen
A;Note: sequence modified after extraction from NCBI backbone
R;Stepan, J.J.; Lau, K.H.W.; Mohan, S.; Singer, F.R.; Baylink, D.J.
Biochem. Biophys. Res. Commun. 168, 792-800, 1990
A;Title: Purification and N-terminal amino acid sequence of the tartrate-resistant acid
A;Reference number: A34638; MUID:90241[illegible]
A;Accession: A34638
A;Molecule type: protein
A;Residues: 22-31 <STE>
A;Experimental source: osteoclastoma
R;Stepan, J.J.; Lau, K.H.W.; Mohan, S.; Kr[illegible]; Baylink, D.J.
Biochem. Biophys. Res. Commun. 165, 1027-1034, 1989
A;Title: Purification and N-terminal sequence [illegible]
A;Reference number: A33781; MUID:90121201
A;Accession: B33781
[illegible]lecule type: protein
[illegible]sidues: 22-33,'X',35-37 <ST3>
[illegible]perimental source: hairy cell leukemia spleen
C;Comment: This enzyme contains a catalytically active binuclear iron [illegible]
C;Genetics:
A;Gene: GDB:ACP5; T5AP
A;Cross-references: GDB:118747; OMIM:171640
A;Map position: 19p13.3-19p13.2
A;Introns: 87/[illegible]
C;Function:
A;Description: catalyzes the hydrolysis of a wide range of phosphate esters
C;Superfamily: tartrate-resistant acid phosphatase; phosphoesterase [illegible]
C;Keywords: blocked amino end; glycoprotein; iron; metalloprotein; phosphoric monoester
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-182,183-325/Product: acid phosphatase ACP5 #status predicted <MAT>
F;22-182/Domain: acid phosphatase ACP5 chain B #status predicted <CHB>
F;27-112/Domain: phosphoesterase core [illegible]
F;183-325/Domain: acid phosphatase ACP5 chain A #status predicted <CHA>
F;33,71,74,242/Binding site: iron (Asp, Asp, Tyr, His) #status predicted
F;71,110,205,240/Binding site: iron (Asp, A[illegible], His, His) #status predicted
F;111,214/Active site: His #status predicted
F;116,147/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;161-219/Disulfide bonds: #status predicted
```

```
F;183/Modified site: blocked amino end (Ile) ([illegible] mature form) (partial) #status exper
F;238/Disulfide bonds: interchain (in [illegible]) #status predicted


  Query Match              54.4%;  Score 1[illegible];  DB 1;  Length 32[illegible];
  Best Local Similarity    68.8%;  Pred. No. 6.[illegible];
  Matches   11;  Conservative    1;  Mismatches    [illegible];  Indels    [illegible];  Gaps    2;

Qy          1 AEHGPI CMV  HPL 13
              ||||||  |        |
Db        212 AEHGPIHELVKQLHPL 227


RESULT    8
JN0787
acid phosphatase (EC 3.1.3.2) 5 precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 24 Feb-1994 #sequence revision [illegible] #text change [illegible]-Mar 2000
C;Accession: JN0787; 155486; 155533
R;Cassady, A.I.; King, A.G.; Cross, N.C.P.; Hume, D.A.
Gene 130, 201-207, 1993
A;Title: Isolation and characterization of the genes encoding mouse and human type 5
A;Reference number: JN0787; MUID:93366175
A;Accession: JN0787
A;Molecule type: DNA
A;Residues: 1-327 <CAS>
A;Cross-references: GB:M99054; NID:g553[illegible]; PIDN:AAA37[illegible].1; PID:g191988
R;Reddy, S.V.; Scarcez, T.; Windle, J.J.; Leach, R.J.; Hundley, J.E.; Chirgwin, J.M.;
J. Bone Miner. Res. 8, 1263-1270, 1993
A;Title: Cloning and characterization of the 5' flanking region of the mouse tartrate
A;Reference number: 155486; MUID:9407882[illegible]
A;Accession: 155486
A;Status: preliminary; translated from [illegible]
A;Molecule type: DNA
A;Residues: 1-48 <RED>
A;Cross-references: GB:M85212; NID:g[illegible]; PIDN:AAA[illegible]; PID:g[illegible]
R;Lacey, D.L.; Erdmann, J.M.; Tan, H.L.
J. Cell. Biochem. 54, 365-371, 1994
A;Title: Interleukin 4 increases type 5 acid phosphatase mRNA expression in murine bo
A;Reference number: 155533; MUID:942[illegible]
A;Accession: 155533
A;Status: preliminary; translated from [illegible]
A;Molecule type: mRNA
A;Residues: 'N',85-193 <LAC>
A;Cross references: GB:S70805; NID:g[illegible]; PIDN:AAB[illegible].1; PID:g546968
C;Genetics:
A;Gene: T5AP; TRAP
A;Introns: 83/3; [illegible]; 247/3
C;Function:
A;Description: catalyzes the hydrolysis of a wide range of phosphate esters
A;Note: also functions as a carrier of iron from the uterine endometrium to the fetus
A;Note: this class of phosphatase has optimal activity at acid pH and is resistant to
C;Superfamily: tartrate-resistant acid phosphatase; phosphoesterase core homology
C;Keywords: blocked amino end; disulfide bond; glycoprotein; iron; metalloprotein; ph
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-184,185-327/Product: acid phosphatase 5 #status predicted <MAT>
F;22-184/Domain: acid phosphatase 5 chain B #status predicted <CHB>
F;29-114/Domain: phosphoesterase core homology <PE>
F;85-[illegible]/Region: [illegible] lysosomal sorting sequence
F;185-327/Domain: acid phosphatase 5 chain A #status predicted <CHA>
F;35,73,76,244/Binding site: iron (Asp, Asp, Tyr, His) #status predicted
F;73,112,207,242/Binding site: iron (Asp, Asn, His, His) #status predicted
F;113,216/Active site: His #status predicted
F;118,149/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;163-221/Disulfide bonds: #status predicted
F;185/Modified site: blocked amino end (Ile) (in mature form) (partial) #status exper
F;219/Disulfide bonds: interchain ([illegible]) #status predicted


  Query Match              54.4%;  Score [illegible];  DB [illegible];  Length 32[illegible];
  Best Local Similarity    68.8%;  Pred. No. 6.[illegible];
  Matches   11;  Conservative    1;  Mismatches    [illegible];  Indels    [illegible];  Gaps    2;
```

```
Qy          1 AEHGPL[illegible]MV [illegible]HH [illegible]13
                [illegible]
Db        214 AEHGPIR[illegible]VENLRI [illegible]

RESULT    9
A41720
acid phosphatase (EC [illegible]) [illegible] precursor [illegible]
C;Species: Rattus norvegicus (Norway rat)
C;Date: 10 Sep [illegible] #sequence_revision [illegible] Sep [illegible]
C;Accession: A417[illegible]
R;Ek Rylander, B.; [illegible] Norgard, M.; N[illegible]
J. Biol. Chem. 26[illegible], [illegible]84 [illegible]
A;Title: Cloning, sequence, and developmental expre[illegible]
A;Reference number: A4172[illegible]; MUID:[illegible]
A;Accession: A4172[illegible]
A;Status: preliminary
A;[illegible]ecule type: mRNA
A;[illegible]sidues: 1-327 [illegible]EKA
A;Cross-references: GB:M7[illegible]; NID:[illegible]; PIDN:AAA[illegible]
C;Superfamily: tartrate-resistant acid phosphatase [illegible]
C;Keywords: phosph[illegible]; mon[illegible]
F;29-114/Domain: phosphoes[illegible]

  Query Match            [illegible]3.4%;  Score 1[illegible]
  Best Local Similarity  [illegible]; Pred. No. [illegible]
  Matches   [illegible];  Conservative [illegible];  Mismatches [illegible]

Qy          1 AEHGPT[illegible]MV [illegible]HH[illegible] 13
                | | || |     |
Db        214 AEHGPTRGVENLRPI [illegible]

RESULT   10
[illegible]4[illegible]34
transketolase (EC [illegible]) [illegible]
C;Species: Alcaligenes eut[illegible]phus
C;Date: 19 May 199[illegible] #sequence_revision [illegible] May [illegible]
C;Accession: [illegible]49934; [illegible]004
R;Schneiderjohann, [illegible]; Kusian, B.; [illegible]
J. Bacteriol. 17[illegible], [illegible]
A;Title: The cbb operon of the fac[illegible]
A;Reference number: A[illegible]34; MUID:[illegible]
A;Accession: [illegible]49934
A;[illegible]tus: preliminary
A;[illegible]ecule type: DNA
A;[illegible]sidues: 1-670 [illegible]
A;Cross-references: GB:M68[illegible]; NID:g4[illegible]; PIDN:AAA[illegible]
R;Kossmann, [illegible]; Klintworth, R.; [illegible]
Gene 85, 247-252, 198[illegible]
A;Title: Sequence ana[illegible]
A;Reference number: [illegible]; MUID:[illegible]
A;Accession: [illegible]0049
A;Molecule type: DNA
A;Residues: 1-77 [illegible]
A;Experimental source: [illegible]
C;Genetics:
A;Genome: plasmid
C;Superfamily: transketolase; thiamin pyrophosphate [illegible]
C;Keywords: transfer[illegible]
F;[illegible]-206/Domain: thiamin pyrophosphate [illegible]

  Query Match            [illegible]
  Best Local Similarity  [illegible]; Pred. No. [illegible]
  Matches   [illegible];  Conservative [illegible];  Mismatches [illegible]

Qy          2 EHGH[illegible]MV[illegible] 10
                [illegible]
Db        51[illegible] ENGPI[illegible]LVI[illegible]
```

```
RES[illegible]
[illegible]N[illegible]
[illegible] (Candida [illegible])
N;Alternate names: [illegible]
C;Species: Candida [illegible]
C;Date: [illegible] #sequence_revision [illegible] #text_change [illegible] Mar [illegible]
C;Accession: [illegible]N[illegible]
R;[illegible] M.; [illegible] Alber[illegible]
Gene [illegible]4, 4[illegible]
A;Title: [illegible] Candida [illegible] phase [illegible]
A;Reference number: [illegible]N[illegible]; MUID:[illegible]
A;A[illegible]N[illegible]
A;M[illegible]type: DNA
A;Re[illegible]
A;[illegible]M[illegible]
A;Note: the [illegible] Candida [illegible]
C;[illegible]
A;Gene: [illegible]
C;[illegible]
C;Keywords: [illegible]
[illegible] #status predicted [illegible]
[illegible] #status predicted MA[illegible]
[illegible] #status predicted

  [illegible] Ma[illegible]
  Be[illegible] Pred. No. [illegible]
  Mat[illegible] Mismatches [illegible]

[illegible]MV[illegible]

[illegible]MN[illegible]

RES[illegible]
A[illegible]
[illegible]
N;Alternate names: [illegible]
C;Species: [illegible]
C;Date: [illegible] Mar [illegible] #sequence_revision [illegible] Mar [illegible] #text_change [illegible] Mar [illegible]
C;A[illegible]
R;Ba[illegible]; A[illegible]; [illegible]ecker, [illegible]; [illegible]
[illegible]
A;Title: [illegible] with [illegible] related protein (SA[illegible]), [illegible] primary structure and [illegible] ated
A;Reference number: A[illegible]; MUID:[illegible]
A;A[illegible]
A;Status: [illegible]
A;Mo[illegible] type: mRNA
A;Re[illegible]
C;Superfamily: [illegible]
C;Keywords: lipoprotein; [illegible]
[illegible] #status predicted
[illegible] phosphate [illegible] #status experimental

  [illegible] Match [illegible]
  Be[illegible] Pred. No. [illegible]
  Mat[illegible] Mismatches [illegible]

[illegible]MV[illegible]A[illegible]

Db [illegible]A[illegible]

[illegible]
[illegible]
[illegible]
C;[illegible]
C;Date: [illegible] #sequence_revision [illegible] #text_change [illegible]
C;A[illegible]
R;[illegible] K.; [illegible] M.; [illegible] W[illegible]; [illegible] M.[illegible]
[illegible] K.; [illegible] Ka[illegible]; A[illegible]
[illegible] M.V[illegible]
```

```
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic pathe
A;Reference number: A82950; MUID:20437337
A;Accession: E83380
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-317 <STO>
A;Cross-references: GB:AE004639; GB:AE004091; NID:g9948125; PIDN:AAG05[illegible] GSPDB:GN00
A;Experimental source: strain PAO1
C;Genetics:
A;Gene: PA2115


  Query Match             51.2%;  Score 41;  DB 2;  Length 317;
  Best Local Similarity   61.5%;  Pred. No. 16;
  Matches    8;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 AEHGPTCMVLHPL 13
                |  || ||:| ||
Db        295 ANVGPECMLLAPL 307


[illegible]LT   14
[illegible]80
probable oxidoreductase PA5327 [imported] - Pseudomonas aeruginosa (strain PAO1)
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 01-Dec-2000
C;Accession: E82980
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrener, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic pathe
A;Reference number: A82950; MUID:20437337
A;Accession: E82980
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-442 <STO>
A;Cross-references: GB:AE004945; GB:AE004091; NID:g9951639; PIDN:AAG08712.1; GSPDB:GN00
A;Experimental source: strain PAO1
C;Genetics:
A;Gene: PA5327
C;Superfamily: L-gulonolactone oxidase


  Query Match             51.2%;  Score 41;  DB 2;  Length 442;
  Best Local Similarity   75.0%;  Pred. No. 14;
  Matches    6;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 AEHGPTCM 8
                ||||| |:
[illegible]       318 AEHGPACL 325


RESULT   15
E86610
axial filament protein [imported] - Chlamydophila pneumoniae (strain J138)
C;Species: Chlamydophila pneumoniae, Chlamydia pneumoniae
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 02-Mar-2001
C;Accession: E86610
R;Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Is
Nucleic Acids Res. 28, 2311-2314, 2000
A;Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.
A;Reference number: A86491; MUID:20330349
A;Accession: E86610
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-515 <STO>
A;Cross-references: GB:BA000008; NID:g8978333; PIDN:BAA99167.1; GSPDB:GN00142
A;Experimental source: strain J138
C;Genetics:
A;Gene: catE
```

```
  Query Match             51.2%;  Score 41;  DB 2;  Length 515;
  Best Local Similarity   54.5%;  Pred. No. [illegible];
  Matches    6;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 EHGPTCMVLHP 12
                ||   |:|:|.
Db        453 EHSHLCLVVHP 463


Search completed: August 22, 2001, 14:2[illegible]4
Job time: 78 sec
```

```
                     GenCore version 4.5
        Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on:          August 22, 2001, 14:29:07 ; Search time [illegible] Seconds
                                              (without alignments)
                                              8.324 Million cell updates/sec

Title:           US-08-659-254-3
Perfect score:   80
Sequence:        1 AEHGPTCMVLHPLA 14

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters:       [illegible]

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       A_Geneseq_0601:*
               1:  /SIDS8/gcgdata/geneseq/[illegible]:*
               2:  /SIDS8/gcgdata/geneseq/[illegible]:*
               3:  /SIDS8/gcgdata/geneseq/[illegible]:*
               4:  /SIDS8/gcgdata/geneseq/[illegible]:*
               5:  /SIDS8/gcgdata/geneseq/[illegible]:*
               6:  /SIDS8/gcgdata/geneseq/[illegible]:*
               7:  /SIDS8/gcgdata/geneseq/[illegible]:*
               8:  /SIDS8/gcgdata/geneseq/[illegible]:*
               9:  /SIDS8/gcgdata/geneseq/[illegible]:*
              10:  /SIDS8/gcgdata/geneseq/[illegible]:*
              11:  /SIDS8/gcgdata/geneseq/[illegible]:*
              12:  /SIDS8/gcgdata/geneseq/[illegible]:*
              13:  /SIDS8/gcgdata/geneseq/[illegible]:*
              14:  /SIDS8/gcgdata/geneseq/[illegible]:*
              15:  /SIDS8/gcgdata/geneseq/[illegible]:*
              16:  /SIDS8/gcgdata/geneseq/[illegible]:*
              17:  /SIDS8/gcgdata/geneseq/[illegible]:*
              18:  /SIDS8/gcgdata/geneseq/[illegible]:*
              19:  /SIDS8/gcgdata/geneseq/[illegible]:*
              20:  /SIDS8/gcgdata/geneseq/[illegible]:*
              21:  /SIDS8/gcgdata/geneseq/[illegible]:*
              22:  /SIDS8/gcgdata/geneseq/[illegible]:*

   Pred. No. is the number of results predicted by chance to have a
   score greater than or equal to the score of the result being printed,
   and is derived by analysis of the total score distribution.

                              SUMMARIES
                  %
Result           Query
  No.   Score    Match Length DB   ID                         Description
-----------------------------------------------------------------------------
   1      80     100.0     14  20  AAW74154                   Mouse steroidogeni
   2      80     100.0     14  22  AAY[illegible]             Mouse tryptic pep
   3      80     100.0    284  17  AAW[illegible]             Human steroidogene
   4      80     100.0    284  20  AAW741[illegible]          Mouse steroidogeni
   5      80     100.0    284  22  AAY[illegible]             Mouse steroidogeni
   6      80     100.0    285  17  AAW[illegible]             Human steroidogene
   7      80     100.0    285  17  AAW[illegible]0814         Human steroidogene
   8      80     100.0    285  17  AAW[illegible]             Human steroidogene
   9      80     100.0    285  17  AAW[illegible]             Human steroidogene
  10      80     100.0    285  17  AAW[illegible]             Human steroidogene
  11      80     100.0    285  22  AAY[illegible]             Human steroidogene
  12      76      95.0    285  17  AAW[illegible]             Human steroidogene
  13      53      66.2      9  22  AAY[illegible]             Steroidogenic peat
  14    43.5      54.4    306  21  AAR[illegible]             Rat tartrate resis
  15    43.5      54.4    332  21  AAR[illegible]             Lung cancer associ
  16      42      52.5    204  20  AAY29[illegible]           Amino acid sequenc
  17      41      51.2    521  20  AAY[illegible]             Chlamydia pneumoni
  18      40      50.0    337  21  AAG27[illegible]           Arabidopsis thalia
  19      40      50.0    217  21  AAG[illegible]             Arabidopsis thalia
  20      40      50.0    220  21  AAG27438                   Arabidopsis thalia
  21      40      50.0    282  21  AAG27[illegible]           Arabidopsis thalia
  22      40      50.0    336  21  AAG49[illegible]           Arabidopsis thalia
  23      40      50.0    374  21  AAG21[illegible]           Arabidopsis thalia
  24      40      50.0    374  21  AAG38[illegible]           Arabidopsis thalia
  25      40      50.0    374  21  AAG40[illegible]           Arabidopsis thalia
  26      40      50.0    400  21  AAG21[illegible]           Arabidopsis thalia
  27      40      50.0    400  21  AAG38[illegible]           Arabidopsis thalia
  28      40      50.0    400  21  AAG40[illegible]           Arabidopsis thalia
  29      40      50.0    404  21  AAG211[illegible]          Arabidopsis thalia
  30      40      50.0    404  21  AAG38[illegible]           Arabidopsis thalia
  31      40      50.0    404  21  AAG40[illegible]           Arabidopsis thalia
  32      40      50.0    423  21  AAG27[illegible]           Arabidopsis thalia
  33      40      50.0    425  21  AAG27[illegible]           Arabidopsis thalia
  34      40      50.0    477  21  AAG49[illegible]           Arabidopsis thalia
  35      40      50.0    479  21  AAG49[illegible]           Arabidopsis thalia
  36      39      48.8    363  21  AAY81[illegible]           Streptococcus pneu
  37      39      48.8    483  21  AAY791[illegible]          Southern copperhea
  38      39      48.8    525  20  AAY[illegible]             Amino acid sequenc
  39      38      47.5    142  21  AAB01[illegible]           Human secreted pro
  40      38      47.5    193  19  AAW98[illegible]           H. pylori GHPO 421
  41      38      47.5    390  21  AAB57[illegible]           Human prostate can
  42      38      47.5    728  14  AAR[illegible]             Hepatocyte growth
  43      38      47.5   1295  20  AAW7[illegible]            Drosophila melanog
  44      38      47.5   1295  22  AAB6[illegible]            Fruit fly patched
  45      38      47.5   1299  16  AAR8[illegible]            Drosophila patched

                              ALIGNMENTS

RESULT 1
AAW74154
ID   AAW74154 standard; Protein; 14 AA.
XX
AC   AAW74154;
XX
DT   [illegible]-MAY-1999  (first entry)
XX
DE   Mouse steroidogenic acute regulatory protein tryptic peptide.
XX
KW   StAR; steroidogenic acute regulatory protein; [illegible] syndrome;
KW   steroid hormone dependent disease; [illegible] congenital LCAH;
KW   lipoid congenital adrenal hyperplasia; [illegible] sexual maturation;
KW   androgen-responsive tumours; hypogonadism; therapy; progesterone level.
XX
OS   Mus sp.
XX
PN   US5872230-A.
XX
PD   16-FEB-1999.
XX
PF   04-NOV-1995;   95US-0538960.
XX
PR   04-NOV-1995;   95US-0538960.
XX
PA   (UYTE ) UNIV TEXAS TECH HEALTH SCI CENT.
XX
PI   Clark BJ,  Stocco DM;
XX
DR   WPI; 1999-166723/14.
XX
PT   New nucleic acid encoding steroidogenic acute regulatory protein -
PT   useful for treatment of steroid hormone-dependent disorders such as
PT   lipoid congenital adrenal hyperplasia, infertility and
PT   androgen responsive tumours
XX
PS   Example 27; Column 29; [illegible]
XX
CC   This sequence is a tryptic peptide fragment [illegible]
CC   regulatory protein (StAR) [illegible]
CC   its products or derivatives may be used as probes [illegible]
CC   hybridisation [illegible]
CC   disorders such as lipoid congenital adrenal hyperplasia (LCAH),
CC   infertility [illegible]
CC   puberty, McCune-Albright syndrome, adrenal hyperplasia [illegible]
CC   hypogonadism [illegible]
CC   progesterone [illegible]
XX
SQ   Sequence 14 AA;

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches 14; Conservative [illegible] Mismatches [illegible]

Qy       1 AFHGPTCMVL[illegible]A 14
             [illegible]
Db       1 aehgptcmvl[illegible]a 14


RESULT 2
AAY72832
ID   AAY72832 standard; peptide; 14 AA.
XX
AC   AAY72832;
XX
DT   31-MAY-2001 [illegible]
XX
DE   Mouse tryptic peptide #23 for analysis [illegible]
XX
KW   Mouse; steroidogenic acute regulatory protein; StAR;
KW   gene therapy; steroid hormone dependent disorder; [illegible]
KW   lipoid congenital adrenal hyperplasia; [illegible]
KW   infertility; androgen responsive tumour; [illegible]
KW   hypogonadotropic hypogonadism; McCune Albright [illegible]
KW   adrenal hyperplasia; [illegible]
XX
OS   Mus sp.
XX
PN   US6194[illegible] B1.
XX
PD   27-FEB-2001.
XX
PF   07-JUN-1996; [illegible]
XX
PR   04-NOV-1995; [illegible]
XX
PA   (UYTE ) UNIV TEXAS [illegible] HEALTH [illegible]
XX
PI   Strauss JM, [illegible]
XX
DR   WPI; 2001-217[illegible]
XX
PT   Novel nucleic acid encoding [illegible]
PT   useful for treating steroid hormone dependent disorder [illegible]
XX
PS   Example [illegible]; Column 29; [illegible]
XX
CC   The invention relates to [illegible]
CC   (StAR) and nucleic [illegible]
CC   are used for regulating steroid [illegible]
CC   proteins are [illegible]
CC   the mitochondria [illegible]
CC   invention are useful in gene therapy [illegible]
CC   hormone dependent disorders [illegible]
CC   adrenal hyperplasia (LCAH) [illegible]
CC   androgen responsive [illegible]
```

```
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible] NA [illegible]
CC   [illegible]
CC   [illegible] AR [illegible]
CC   [illegible]
CC   [illegible] StAR protein.
XX
SQ   Sequence 14 AA;

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches [illegible] Mismatches [illegible]

Qy       1 AF[illegible]MV[illegible] 14

Db       1 [illegible] 14


RESULT 3
AAW[illegible]
ID   AAW[illegible] standard; peptide; 14 AA.
XX
AC   AAW[illegible]
XX
DT   [illegible] MAR [illegible]
XX
DE   [illegible]
XX
KW   [illegible] StAR [illegible]
KW   [illegible]
KW   [illegible]
KW   [illegible]
XX
OS   Homo sapiens.
XX
PN   WO[illegible]
XX
PD   [illegible] SEP [illegible]
XX
PF   [illegible] MAR [illegible] W[illegible]
XX
PR   [illegible] MAR [illegible]
XX
PA   ([illegible]) UNIV CALIFORNIA
PA   (UYPE ) UNIV PENNSYLVANIA
XX
PI   [illegible] M, [illegible] W, [illegible]
XX
DR   WPI; [illegible]
XX
PT   [illegible]
PT   [illegible]
PT   [illegible] adrenal hyperplasia
XX
PS   [illegible]
XX
CC   The present sequence [illegible]
CC   [illegible] (StAR) [illegible] (StAR)
CC   [illegible] (LCAH) [illegible]
CC   [illegible] StAR [illegible]
CC   [illegible] predisposition [illegible] StAR associated
CC   [illegible] StAR [illegible]
CC   [illegible] therapy which can also be
CC   [illegible]
XX
SQ   Sequence [illegible] AA;

  Query Match [illegible]
  Best Local Similarity   100.0%;  Pred. No. [illegible]e-05;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 AEHGPTCMVLHPLA 14
            ||||||||||||||
Db      218 aehqptcmvlhpla 231


RESULT    4
AAW74153
ID   AAW74153 standard; Protein; 284 AA
XX
AC   AAW74153;
XX
DT   05-MAY-1999  (first entry)
XX
DE   Mouse steroidogenic acute regulatory protein.
XX
KW   StAR; steroidogenic acute regulatory protein; McCune-Albright syndrome;
KW   steroid hormone dependent disorder; adrenal hypoplasia congenita; LCAH;
KW   lipoid congenital adrenal hyperplasia; infertility; sexual maturation;
[illegible]   androgen responsive tumours; hypogonadism; therapy; progesterone level.
[illegible]
[illegible]   Mus sp.
XX
PN   US5872230-A.
XX
PD   16-FEB-1999.
XX
PF   04-NOV-1995;   95US-0538960.
XX
PR   04-NOV-1995;   95US-0538960.
XX
PA   (UYTE-) UNIV TEXAS TECH HEALTH SC CENT.
XX
PI   Clark BJ,  Stocco DM;
XX
DR   WPI; 1999-166720/14.
DR   N-PSDB; AAX18169, AAX18170.
XX
PT   New nucleic acid encoding steroidogenic acute regulatory protein
PT   useful for treatment of steroid hormone dependent disorders such as
PT   lipoid congenital adrenal hyperplasia, infertility and
PT   androgen-responsive tumours
XX
PS   Claim 1; Column 37-38; 25pp; English.
XX
CC   This sequence is the steroidogenic acute regulatory protein (StAR)
CC   of the invention. The StAR coding sequence and its products or
[illegible]   derivatives may be used as probes and primers for PCR hybridisation
[illegible]   studies and in the treatment of steroid hormone dependent disorders such
[illegible]   as lipoid congenital adrenal hyperplasia (LCAH), infertility, sexual
CC   maturation, androgen-responsive tumours, precocious puberty,
CC   McCune-Albright syndrome, adrenal-hypoplasia congenita, and hypogonadism.
CC   Monitoring levels of StAR may also indicate reduced progesterone levels
CC   and so a risk of spontaneous abortion.
XX
SQ   Sequence   284 AA;


  Query Match             100.0%;  Score [illegible];  DB 20;  Length 284;
  Best Local Similarity   100.0%;  Pred. No. [illegible]e-05;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        1 AEHGPTCMVLHPLA 14
            ||||||||||||||
Db      217 aehqptcmvlhpla 230


RESULT    5
AAY72831
ID   AAY72831 standard; Protein; 284 AA.
XX
```

```
AC   AAY72831;
XX
DT   31-MAY-2001  (first entry)
XX
DE   Mouse steroidogenic acute regulatory (StAR) protein.
XX
KW   Mouse; steroidogenic acute regulatory protein; StAR;
KW   gene therapy; steroid hormone-dependent disorder; LCAH;
KW   lipoid congenital adrenal hyperplasia; sexual maturation;
KW   infertility; androgen-responsive tumour; precocious puberty;
KW   hypogonadotropic hypogonadism; McCune-Albright syndrome;
KW   adrenal-hypoplasia congenita; cytostatic.
XX
OS   Mus sp.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..47
FT                   /label= Signal_peptide
FT   Cleavage-site   39..40
FT                   /note= "Site for mitochondrial proteases,
FT                   matrix processing protease and mitochondrial
FT                   intermediate processing peptide"
FT   Protein         48..284
FT                   /label= Mature_steroidogenic acute regulatory_protein
FT   Cleavage-site   49..50
FT                   /note= "Site for mitochondrial proteases,
FT                   matrix processing protease and mitochondrial
FT                   intermediate processing peptide"
FT   Modified-site   59..60
FT                   /note= "PKA phosphorylation site"
FT   Modified site   68..69
FT                   /note= "CK phosphorylation site"
FT   Region          85..96
FT                   /label= Tryptic_peptide #2
FT   Region          88..98
FT                   /note= "This sequence is used to raise antibodies
FT                   against StAR protein"
FT   Region          193..211
FT                   /label= Tryptic_peptide #4
FT   Modified-site   195
FT                   /note= "PKA phosphorylation site"
FT   Region          217..230
FT                   /label= Tryptic_peptide #5
FT   Modified-site   233..234
FT                   /note= "P34CDC2 kinase (?) phosphorylation site"
FT   Modified-site   238
FT                   /note= "PKC phosphorylation site"
XX
PN   US6194555-B1.
XX
PD   27-FEB-2001.
XX
PF   07-JUN-1996;   96US-0659254.
XX
PR   04-NOV-1995;   95US-0538960.
XX
PA   (UYTE-) UNIV TEXAS TECH HEALTH SCI CENT.
XX
PI   Stocco DM,  Clark BJ;
XX
DR   WPI; 2001-217436/22.
DR   N-PSDB; AAD62901, AAD62902.
XX
PT   Novel nucleic acid encoding steroidogenic acute regulatory protein,
PT   useful for treating steroid hormone dependent disorders
XX
PS   Example 2; Fig 2; 41pp; English.
XX
CC   The invention relates to steroidogenic acute regulatory protein
CC   (StAR) and nucleic acid molecules encoding them. StAR proteins
CC   are used for regulating steroidogenesis. More particularly these
CC   proteins are involved in regulating cholesterol transport into
CC   the mitochondria of a cell. The nucleic acid molecules of the
CC   invention are useful in gene therapy [illegible]
CC   hormone dependent [illegible]
CC   adrenal hyperplasia (CAH) [illegible]
CC   androgen [illegible]
CC   syndrome, adrenal hypoplasia [illegible]
CC   hypogonadism. [illegible]
CC   of the steroidogenic [illegible]
CC   The present [illegible]
CC   protein (StAR)
XX
SQ   Sequence [illegible]

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches [illegible]

[illegible] 1 AHEGPL[illegible]A [illegible]
[illegible]
[illegible] 2.7 [illegible]

RESULT 6
AAW[illegible]
ID   AAW[illegible] standard; [illegible] AA.
XX
AC   AAW[illegible]
XX
DT   21-MAY-1997 [illegible]
XX
DE   Human steroidogenesis [illegible]
XX
KW   Human; steroidogenesis [illegible]
KW   mutation detection; [illegible]
KW   lipoid adrenal hyperplasia; treatment; [illegible]
KW   replacement therapy; [illegible]
XX
OS   Homo sapiens.
XX
FH   Key           [illegible]
FT   Misc-difference [illegible]
FT                 /note= "wild type [illegible]"
XX
PN   WO9629[illegible] A1.
XX
PD   26-SEP-1996.
XX
[illegible] 22-MAR-1996; [illegible]
XX
PR   23-MAR-1995; [illegible]
XX
PA   (REGC ) UNIV [illegible]
PA   (GYPE-) UNIV PENNSYLVANIA.
XX
PI   [illegible] Miller W [illegible]
XX
DR   WPI; [illegible] 44[illegible]
DR   N-PSDB; AAT6[illegible]
XX
PT   Isolated human [illegible]
PT   for detection [illegible]
PT   lipoid adrenal hyperplasia [illegible]
XX
PS   Claim [illegible]
XX
CC   The present [illegible]
CC   protein (hStAR) [illegible]
CC   associated mutant [illegible]
CC   mutation [illegible]
CC   associated genetic defect [illegible]
CC   can be treated [illegible]
CC   also be used to prevent [illegible]
XX
```

```
[illegible]

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches [illegible]

[illegible] AHEGPL[illegible]

[illegible]

RESULT [illegible]
AAW[illegible]
ID   AAW[illegible] standard; [illegible] AA.
XX
AC   AAW[illegible]
XX
DT   [illegible] MAY [illegible]
XX
DE   [illegible]
XX
KW   [illegible] (StAR) [illegible]
KW   [illegible]
KW   [illegible]
KW   [illegible]
XX
OS   [illegible]
XX
FH   [illegible]
FT   [illegible] "[illegible] kinase"
FT   [illegible] phosph[illegible]"
FT   [illegible] kinase A
FT   [illegible]"
FT   [illegible] kinase A
FT   [illegible]"
FT   [illegible] kinase A
FT   [illegible]"
FT   [illegible] kinase A
FT   [illegible]"
FT   [illegible] kinase A
FT   [illegible]"
XX
PN   [illegible]
XX
PD   [illegible]
XX
[illegible] MAR [illegible]
XX
[illegible] MAR [illegible]
XX
PA   [illegible]
PA   [illegible]
XX
PI   [illegible]
XX
DR   [illegible]
DR   [illegible]
XX
PT   [illegible]
PT   [illegible]
PT   lipoid adrenal hyperplasia
XX
PS   Claim 1; Pages 62-63; 89pp; English.
XX
CC   The present sequence is the human steroidogenesis acute
CC   regulatory protein (hStAR). The hStAR gene can be analysed for
CC   mutations to detect (e.g. prenatally) genetic defects associated
CC   with congenital lipoid adrenal hyperplasia (CAH), or its
CC   transmission to children. CAH can be treated by protein or gene
CC   replacement therapy, which can also be used to prevent or treat
CC   hypercholesterolaemia.
CC   A human adrenal cortex cDNA library was screened with a mouse StAR
CC   probe to isolate a 1.6 kb insert, including an ORF for a 285
CC   residue protein. When it was cloned into pSPORT and expressed in
CC   COS-1 cells cotransfected with pP450scc and pADX, it increased the
CC   level of pregnenolone synthesis from cholesterol or
CC   20-alpha-hydroxycholesterol.
XX
SQ   Sequence   285 AA;

 uery Match              100.0%;  Score [illegible];  DB 17;  Length 285;
 st Local Similarity   100.0%;  Pred. No. [illegible];
 tches   14;  Conservative    0;  Mismatches    0;  Indels    [illegible];  Gaps    [illegible];

Qy        1 AEHGPTCMVLHPLA 14
              ||||||||||||||
Db      218 aehgptcmvlhpla 231


RESULT   8
AAW13233
ID   AAW13233 standard; Protein; 285 AA.
XX
AC   AAW13233;
XX
DT   21-MAY-1997  (first entry)
XX
DE   Human steroidogenesis acute regulatory protein mutant Glu169Gly.
XX
KW   Human; steroidogenesis; acute regulatory protein; hStAR; analysis;
KW   mutation; detection; prenatal; genetic defect; congenital; protein;
KW   lipoid adrenal hyperplasia; treatment; prevention; gene;
KW   replacement therapy; hypercholesterolaemia; mutant.
XX
OS   Homo sapiens.
XX
FH   Key                 Location/Qualifiers
FT   Misc-difference 169
FT                       /note= "wild type Glu substituted with Gly"

     WO9629338-A1.
XX
PD   26-SEP-1996.
XX
PF   22-MAR-1996;   96WO-US03896.
XX
PR   23-MAR-1995;   95US-0410540.
XX
PA   (REGC ) UNIV CALIFORNIA.
PA   (UYPE-) UNIV PENNSYLVANIA.
XX
PI   Lin D,  Miller WL,  Strauss JF;
XX
DR   WPI; 1996-443130/44.
DR   N-PSDB; AAT62018.
XX
PT   Isolated human steroidogenesis acute regulatory protein gene - used
PT   for detection of mutation(s) of this gene that cause congenital
PT   lipoid adrenal hyperplasia
XX
PS   Claim 13; Page  ; 89pp; English.
XX
```

```
CC   The present sequence is the human steroidogenesis acute regulatory
CC   protein (hStAR), congenital lipoid adrenal hyperplasia (CAH)
CC   associated mutant, Glu169Gly. Analysing the hStAR gene for such a
CC   mutation is useful in the detection (e.g. prenatally) of CAH
CC   associated genetic defects, or its transmission to children. CAH
CC   can be treated by protein or gene replacement therapy, which can
CC   also be used to prevent or treat hypercholesterolaemia.
XX
SQ   Sequence   285 AA;

  Query Match             100.0%;  Score [illegible];  DB 17;  Length 285;
  Best Local Similarity   100.0%;  Pred. No. [illegible];
  Matches   14;  Conservative    0;  Mismatches    [illegible];  Indels    [illegible];  Gaps    [illegible];

Qy        1 AEHGPTCMVLHPLA 14
              ||||||||||||||
Db      218 aehgptcmvlhpla 231


RESULT   9
AAW13234
ID   AAW13234 standard; Protein; 285 AA.
XX
AC   AAW13234;
XX
DT   21-MAY-1997  (first entry)
XX
DE   Human steroidogenesis acute regulatory protein mutant Arg182Leu.
XX
KW   Human; steroidogenesis; acute regulatory protein; hStAR; analysis;
KW   mutation; detection; prenatal; genetic defect; congenital; protein;
KW   lipoid adrenal hyperplasia; treatment; prevention; gene;
KW   replacement therapy; hypercholesterolaemia; mutant.
XX
OS   Homo sapiens.
XX
FH   Key                 Location/Qualifiers
FT   Misc-difference 182
FT                       /note= "wild type Arg substituted with Leu"
XX
PN   WO9629338-A1.
XX
PD   26-SEP-1996.
XX
PF   22-MAR-1996;   96WO-US03896.
XX
PR   23-MAR-1995;   95US-0410540.
XX
PA   (REGC ) UNIV CALIFORNIA.
PA   (UYPE-) UNIV PENNSYLVANIA.
XX
PI   Lin D,  Miller WL,  Strauss JF;
XX
DR   WPI; 1996-443130/44.
DR   N-PSDB; AAT62019.
XX
PT   Isolated human steroidogenesis acute regulatory protein gene - used
PT   for detection of mutation(s) of this gene that cause congenital
PT   lipoid adrenal hyperplasia
XX
PS   Claim 13; Page  ; 89pp; English.
XX
CC   The present sequence is the human steroidogenesis acute regulatory
CC   protein (hStAR), congenital lipoid adrenal hyperplasia (CAH)
CC   associated mutant, Arg182Leu. Analysing the hStAR gene for such a
CC   mutation is useful in the detection (e.g. prenatally) of CAH
CC   associated genetic defects, or its transmission to children. CAH
CC   can be treated by protein or gene replacement therapy, which can
CC   also be used to prevent or treat hypercholesterolaemia.
XX
SQ   Sequence   285 AA;
```

```
  Query Match             [illegible]
  Best Local Similarity   [illegible]
  Matches   [illegible]  Conservative [illegible]  Mismatches [illegible]

Qy          1 AEHGPPMV[illegible]LA [illegible]
                [illegible]
Db        218 aehgpt[illegible]

RESULT 10
AAW13235
ID   AAW13235 standard; Protein; [illegible] AA.
XX
AC   AAW13235;
XX
DT   21 MAY 1997  (first entry)
XX
DE   Human steroidogenesis acute regulatory [illegible]
XX
KW   Human; steroidogenesis acute regulatory [illegible]
KW   mutation; detection; [illegible]
KW   lipoid adrenal hyperplasia; treatment; [illegible]
KW   replacement therapy; [illegible]
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference [illegible]
FT                   /note= "[illegible]"
XX
PN   WO96[illegible]-A1.
XX
PD   26-SEP-1996.
XX
PF   22-MAR-1996; [illegible]
XX
PR   23-MAR-1995; [illegible]
XX
PA   (REGC ) UNIV CALIFORNIA.
PA   (UYPE-) UNIV PENNSYLVANIA.
XX
PI   Lin D,  Miller WL,  [illegible]
XX
DR   WPI; 1996-44[illegible]
DR   N-PSDB; AAT[illegible]
XX
PT   Isolated human steroidogenesis acute regulatory [illegible]
PT   for detection of mutation(s) of [illegible]
PT   lipoid adrenal hyperplasia [illegible]
XX
PS   Claim [illegible]; Page [illegible]; [illegible]pp; [illegible]
XX
CC   The present sequence [illegible] the human [illegible]
CC   protein (hSTAR) [illegible] lipoid adrenal hyperplasia [illegible]
CC   associated mutant [illegible]
CC   mutation is used in the detection [illegible]
CC   associated genetic defect [illegible]
CC   can be treated by [illegible]
CC   also be used to [illegible]
XX
SQ   Sequence   [illegible] AA;

  Query Match             [illegible]
  Best Local Similarity   [illegible]
  Matches   [illegible]  Conservative [illegible]  Mismatches [illegible]

Qy          1 AE[illegible]PCMV[illegible] 14
                [illegible]
Db        218 aehgpt[illegible]
```

```
RESULT [illegible]
AAY[illegible]
ID   AAY[illegible] standard; Protein; [illegible] AA.
XX
AC   AAY[illegible]
XX
DT   [illegible] MAY [illegible] (first entry)
XX
DE   Human steroidogenic acute regulatory (StAR) proteins.
XX
KW   Human; steroidogenic acute regulatory protein; StAR;
KW   gene therapy; steroid hormone dependent disorder; CAH;
KW   lipoid congenital adrenal hyperplasia; sexual maturation;
KW   [illegible] puberty;
KW   hypogonadotropic hypogonadism; McCune Albright syndrome;
KW   [illegible]
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference [illegible]
FT                   [illegible]
FT                   /note= "the translation except [illegible] the sequence shown
FT                   [illegible]"
FT   Modified-site   [illegible]
FT                   /note= "[illegible]"
FT   Modified-site   [illegible]
FT                   /note= "[illegible]"
FT   Modified-site   [illegible]
FT                   /note= "[illegible]"
FT   Modified-site   [illegible]
FT                   /note= "[illegible]"
FT   Modified-site   [illegible]
FT                   /note= "[illegible]"
FT   Modified-site   [illegible]
FT                   /note= "[illegible]"
XX
PN   [illegible]
XX
PD   [illegible]
XX
PF   [illegible]
XX
PR   [illegible]
XX
PA   [illegible] HEALTH [illegible]
XX
PI   [illegible]
XX
DR   WPI; [illegible]
DR   N-PSDB; [illegible]
XX
PT   [illegible] steroidogenic acute regulatory proteins [illegible]
PT   use [illegible] treating steroid hormone dependent disorders
XX
PS   Example [illegible]
XX
CC   The present invention relates to steroidogenic acute regulatory proteins
CC   [illegible] (StAR) proteins
CC   are used for regulating steroidogenesis. More particularly, these
CC   proteins are [illegible] transport [illegible]
CC   the mitochondria [illegible]
CC   [illegible] therapy for the treatment of [illegible]
CC   hormone dependent [illegible] particularly lipoid [illegible]
CC   adrenal [illegible] sexual maturation [illegible]
CC   and [illegible] McCune Albright
CC   [illegible]
CC   [illegible] preparation [illegible]
CC   [illegible]
CC   The present sequence [illegible] steroidogenic acute regulatory
CC   protein [illegible]
XX
SQ   Sequence   [illegible]
```

```
  Query Match             100.0%;  Score 80;  DB 22;  Length 285;
  Best Local Similarity   100.0%;  Pred. No. [illegible];
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AEHGPTCMVLHPLA 14
              ||||||||||||||
Db        218 aehqptcmvlhpla 231


RESULT 12
AAW13236
ID   AAW13236 standard; Protein; 285 AA
XX
AC   AAW13236;
XX
DT   21-MAY-1997  (first entry)
XX
DE   Human steroidogenesis acute regulatory protein mutant Ala218Val.
XX
KW   Human; steroidogenesis; acute regulatory protein; hStAR; analysis;
     mutation; detection; prenatal; genetic defect; congenital; protein;
     lipoid adrenal hyperplasia; treatment; prevention; gene;
KW   replacement therapy; hypercholesterolaemia; mutant.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 218
FT                   /note= "wild type A is substituted with V."
XX
PN   WO9629338-A1.
XX
PD   26-SEP-1996.
XX
PF   22-MAR-1996;   96WO-US03896.
XX
PR   23-MAR-1995;   95US-0410540.
XX
PA   (REGC ) UNIV CALIFORNIA.
PA   (UYPE-) UNIV PENNSYLVANIA.
XX
PI   Lin D,  Miller WL,  Strauss JF;
XX
DR   WPI; 1996-443130/44.
DR   N-PSDB; AAT62021.
XX
PT   Isolated human steroidogenesis acute regulatory protein gene - used
PT   for detection of mutation(s) of this gene that cause congenital
     lipoid adrenal hyperplasia
XX
PS   Claim 13; Page -; 89pp; English.
XX
CC   The present sequence is the human steroidogenesis acute regulatory
CC   protein (hStAR), congenital lipoid adrenal hyperplasia (CAH)
CC   associated mutant, Ala218Val. Analysing the hStAR gene for such a
CC   mutation is useful in the detection (e.g. prenatally) of CAH
CC   associated genetic defects, or its transmission to children. CAH
CC   can be treated by protein or gene replacement therapy, which can
CC   also be used to prevent or treat hypercholesterolaemia.
XX
SQ   Sequence   285 AA;


  Query Match             95.0%;  Score [illegible];  DB 17;  Length 285;
  Best Local Similarity   100.0%;  Pred. No. 8.5e-05;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 EHGPTCMVLHPLA 14
              |||||||||||||
Db        219 ehqptcmvlhpla 231
```

```
RESULT 13
AAY72836
ID   AAY72836 standard; peptide; 9 AA.
XX
AC   AAY72836;
XX
DT   31-MAY-2001  (first entry)
XX
DE   Steroidogenic acute regulatory (StAR) peptide #2.
XX
KW   Steroidogenic acute regulatory protein; StAR;
KW   gene therapy, steroid hormone dependent disorder; LCAH;
KW   lipoid congenital adrenal hyperplasia; sexual maturation;
KW   infertility; androgen responsive tumour; precocious puberty;
KW   hypogonadotropic hypogonadism; McCune-Albright syndrome;
KW   adrenal-hypoplasia congenita; cytostatic.
XX
OS   Unidentified.
XX
PN   US6194555-B1.
XX
PD   27-FEB-2001.
XX
PF   07-JUN-1996;   96US-0659254.
XX
PR   04-NOV-1995;   95US-0538960.
XX
PA   (UYTE ) UNIV TEXAS TECH HEALTH SCI CENT.
XX
PI   Stocco DM,  Clark BJ;
XX
DR   WPI; 2001-217936/22.
XX
PT   Novel nucleic acid encoding steroidogenic acute regulatory protein,
PT   useful for treating steroid hormone dependent disorders
XX
PS   Disclosure; Column 12; 41pp; English.
XX
CC   The invention relates to steroidogenic acute regulatory protein
CC   (StAR) and nucleic acid molecules encoding them. StAR proteins
CC   are used for regulating steroidogenesis. More particularly these
CC   proteins are involved in regulating cholesterol transport into
CC   the mitochondria of a cell. The nucleic acid molecules of the
CC   invention are useful in gene therapy for the treatment of steroid
CC   hormone-dependent disorders, particularly, lipoid congenital
CC   adrenal hyperplasia (LCAH), infertility, sexual maturation,
CC   androgen-responsive tumours, precocious puberty, McCune-Albright
CC   syndrome, adrenal hypoplasia congenita or hypogonadotropic
CC   hypogonadism. They are also useful for recombinant preparation
CC   of the steroidogenic acute regulatory protein.
CC   The present sequence is steroidogenic acute regulatory (StAR)
CC   peptide. This peptide is useful for designing oligonucleotides for
CC   screening and identifying antigenic determinants of StAR protein.
XX
SQ   Sequence   9 AA;


  Query Match             66.2%;  Score [illegible];  DB 22;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. [illegible];
  Matches    9;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 AEHGPTCMV 9
              |||||||||
Db          1 aehqptcmv 9


RESULT 14
AAB13356
ID   AAB13356 standard; protein; 30[illegible] AA
XX
AC   AAB13356;
XX
```

```
[illegible]  21 NOV 2000  ([illegible])
XX
[illegible]  Rat tartrate [illegible] purple acid [illegible]
XX
KW  Rat; tartrate [illegible] purple acid [illegible]
KW  purple acid phosphatase; TAP [illegible]
KW  bone metabolism [illegible]
KW  protein [illegible]
XX
[illegible]  Rattus sp.
XX
PN  WO200[illegible]A[illegible]
XX
[illegible]  2[illegible] JUN 200[illegible]
XX
[illegible]  2[illegible] NOV 1999; [illegible]WO[illegible]
XX
[illegible]  18 DEC 1998; [illegible]E [illegible]
[illegible]  (PHAA ) PHARMA[illegible]
XX
[illegible]  Uppenberg [illegible]
XX
[illegible]  WPI; 200[illegible] 44[illegible]
XX
[illegible]  Crystalline form of tartrate resistant [illegible]
[illegible]  activated by cleavage [illegible] useful for treatment [illegible]
[illegible]  conditions [illegible]
XX
[illegible]  Disclosure; [illegible]
XX
[illegible]  The present sequence [illegible] rat tartrate resistant [illegible]
[illegible]  (TRAP). TRAP [illegible]
[illegible]  phosphatases (PAPs), which are characterised [illegible]
[illegible]  at the active site. TRAP is expressed [illegible]
[illegible]  cells (osteoclasts), where it can mediate [illegible]
[illegible]  matrix proteins such as osteopontin [illegible]
[illegible]  form of activated TRAP was prepared by [illegible]
[illegible]  protease and then crystallising the product [illegible]
[illegible]  used for structure based drug design for [illegible]
[illegible]  and inhibitors of TRAP activity. These [illegible]
[illegible]  of diseases or degenerative conditions [illegible]
[illegible]  resorption, such as [illegible]
[illegible]  osteoporosis
XX
[illegible]  Sequence [illegible] AA;

  Query Match  [illegible]
  Best Local Similarity [illegible]
  Matches [illegible]  Mismatches [illegible]

Qy  [illegible] AEHGPT[illegible]MV [illegible] HPL [illegible]
           |  |            
Db  [illegible] adgaptreck[illegible] rpl[illegible]


RESULT [illegible]
AAB[illegible]
ID  AAB[illegible] standard; [illegible] AA.
XX
AC  AAB[illegible]
XX
DT  14 MAR 2001  ([illegible])
XX
DE  Lung cancer associated [illegible] peptide [illegible]
XX
KW  Human; lung cancer associated protein; [illegible]
KW  [illegible]
KW  gastrointestinal; [illegible]
KW  antibacterial; [illegible]
KW  proliferative disorder; wound [illegible]
XX
[illegible]
XX
PN  W[illegible] A[illegible]
XX
[illegible]
XX
[illegible] MAR [illegible]
XX
[illegible] MAR [illegible]
XX
PA  [illegible] MAN [illegible] MED [illegible] N
PA  [illegible] TEN [illegible] A
XX
[illegible]
XX
[illegible]  W[illegible]
[illegible]  N[illegible] AA[illegible]
XX
[illegible]  [illegible] patients [illegible]
[illegible]  [illegible] treatment [illegible] prevent [illegible]
[illegible]  [illegible]
XX
[illegible]  [illegible] English.
XX
[illegible]  [illegible] sequences AAF[illegible] AAF[illegible] encode human lung cancer
[illegible]  [illegible] AAB[illegible] AAB[illegible]
[illegible]  associated proteins and polynucleotide sequences. Their expression and
[illegible]  [illegible]
[illegible]  [illegible] inflammatory [illegible] gastrointestinal
[illegible]  [illegible]
[illegible]  [illegible] specific for the
[illegible]  [illegible] isolated
[illegible]  polynucleotide sequences may be used for [illegible] lung cancer
[illegible]  [illegible] other
[illegible]  [illegible] purposes. The proteins may be used to treat
[illegible]  [illegible]
[illegible]  gastrointestinal [illegible] and proliferative
[illegible]  disorders. The [illegible] treatment of wounds and
[illegible]  [illegible] AAB[illegible] AAB[illegible] and
[illegible]  peptide AAB[illegible]
[illegible]  [illegible] and protein
[illegible]  sequence
XX
SQ  Sequence [illegible] AA;

  Query Match  [illegible]
  Best Local Similarity [illegible] Pred. No. [illegible]
  Matches [illegible]  Mismatches [illegible]  Indels [illegible]  Gaps [illegible]

Qy  [illegible] A[illegible]MV [illegible]

Db  [illegible]


Search completed: A[illegible]
[illegible]
```

```
                         GenCore version 4.5
             Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on:          August 22, 2001, 14:24:06 ; Search time 55.8 Seconds
                                             (without alignments)
                                             574.093 Million cell updates/sec

Title:           US-08-659-254-2
Perfect score:   1429
Sequence:        1 MFLATFKLCAGSSYRHMRNM..........QVLSQIGFEANEKKELA 276

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters:      93435

   mum DB seq length: 0
   mum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       SwissProt_39:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                              SUMMARIES
                  %
Result          Query
  No.   Score   Match Length DB  ID                            Description
-------------------------------------------------------------------------------
    1    1429   100.0    284  1  STAR_MOUSE                    P51557 mus musculu
    2    1366    95.6    284  1  STAR_RAT                      P97826 rattus norv
    3    1325    92.7    284  1  STAR_MESAU                    P79114 mesocricetu
    4  1262.5    88.3    285  1  STAR_HORSE                    O46689 equus cabal
    5  1255.5    87.9    285  1  STAR_PIG                      Q28990 sus scrofa
    6  1252.5    87.6    285  1  STAR_HUMAN                    P49675 homo sapien
    7  1214.5    85.0    285  1  STAR_BOVIN                    Q28918 bos taurus
    8     534    37.4    121  1  STAR_SHEEP                    P79245 ovis aries
    9     374    26.2    445  1  ML64_HUMAN                    Q14849 homo sapien
   10     367    25.7    446  1  ML64_MOUSE                    Q61542 mus musculu
   11     181    12.7    241  1  YX70_CAEEL                    Q17??? caenorhabdi
   12      93     6.5    295  1  FBX2_HUMAN                    Q9UK22 homo sapien
   13      92     6.4    527  1  RF3_HAEIN                     P43928 haemophilus
   14    89.5     6.3    422  1  TERL_BPSPP                    P54308 bacteriopha
   15    89.5     6.3    554  1  HYES_MOUSE                    P34914 mus musculu
   16      89     6.2    268  1  Y064_MYCTU                    O53194 mycobacteri
   17      88     6.2    404  1  POTG_ECOLI                    P31134 escherichia
   18    87.5     6.1   1047  1  HIRA_DROME                    O17468 drosophila
   19    86.5     6.1    636  1  GYRB_THEMA                    P77993 thermotoga
   20    86.5     6.1    994  1  ATC1_CHICK                    P13585 g sarcoplas
   21    85.5     6.0    546  1  TCPQ_SCHPO                    P78921 schizosacch
   22      85     5.9    554  1  HYES_RAT                      P80299 rattus norv
   23      85     5.9   3433  1  UTRO_HUMAN                    P46939 homo sapien
   24    84.5     5.9    338  1  CTE2_HUMAN                    O00154 homo sapien
   25      84     5.9    605  1  VE1_HPV3?                     P50??? human papi
   26    83.5     5.8    338  1  CTE2_RAT                      Q64559 rattus norv
   27      83     5.8    197  1  RB12_RAT                      P35289 rattus norv
   28    82.5     5.8    574  1  ERG1_HUMAN                    Q14534 homo sapien
   29      82     5.7    474  1  KCC4_RAT                      P13234 rattus norv
   30      82     5.7    701  1  EFG_MYCTU                     O53??? mycobacteri
   31      82     5.7    950  1  DC12_DROME                    P18??? drosophila
   32      82     5.7   1123  1  DC11_DROME                    P18169 drosophila
   33      82     5.7   1589  1  DC13_DROME                    P18??? drosophila
```

```
   34    81.5     5.7    257  1  BLA2_BACSP                    P10425 bacillus sp
   35      81     5.7    389  1  IHFL_MYCTU                    Q10625 mycobacteri
   36      81     5.7   1207  1  EGF_HUMAN                     P01133 homo sapien
   37    80.5     5.6    994  1  ATC1_RAT                      Q64578 r sarcoplas
   38      80     5.6   1008  1  SN14_YEAST                    P36048 saccharomyc
   39    79.5     5.6    972  1  SYL_MYCTU                     Q50062 mycobacteri
   40    79.5     5.6    997  1  ATC2_CANFA                    Q00?04 s sarcoplas
   41    79.5     5.6    997  1  ATC2_FELCA                    Q00779 t sarcoplas
   42    79.5     5.6   1001  1  ATC1_RABIT                    P04191 s sarcoplas
   43    79.5     5.6   1041  1  ATC2_CHICK                    Q03669 g sarcoplas
   44    79.5     5.6   1042  1  ATC2_HUMAN                    P16615 h sarcoplas
   45    79.5     5.6   1042  1  ATC2_PIG                      P11607 s sarcoplas


                              ALIGNMENTS


RESULT   1
STAR_MOUSE
ID   STAR_MOUSE      STANDARD;      PRT;    284 AA.
AC   P51557;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   30 MAY 2000 (Rel. 39, Last annotation update)
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN PRECURSOR (LUTEINIZING HORMONE-
DE   INDUCED PROTEIN).
GN   STAR.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID 10090;
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX   MEDLINE 95050616; PubMed 7961770;
RA   Clark B.J., Wells J., King S.R., Stocco D.M.;
RT   "The purification, cloning, and expression of a novel luteinizing
RT   hormone-induced mitochondrial protein in MA-10 mouse Leydig tumor
RT   cells. Characterization of the steroidogenic acute regulatory protein
RT   (StAR).";
RL   J. Biol. Chem. 269:28314-28322(1994).
CC   -!- FUNCTION: PLAYS A KEY ROLE IN STEROID HORMONE SYNTHESIS BY
CC       ENHANCING THE METABOLISM OF CHOLESTEROL INTO PREGNENOLONE.
CC       TRANSPORTER THAT BINDS TO AND TRANSFER CHOLESTEROL THROUGH THE
CC       INTERMEMBRANE SPACE OF THE MITOCHONDRION (PROBABLE).
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL.
CC   -!- SIMILARITY: BELONGS TO THE STAR FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its  content  is  in  no  way
CC   modified and this statement is not  removed.  Usage  by  and for  commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; L36062; AAB94783.1; -.
DR   MGD; MGI:102760; Star.
DR   InterPro; IPR000799; -.
DR   InterPro; IPR002913; -.
DR   Pfam; PF01852; START; 1.
DR   PRINTS; PR00978; STARPROTEIN.
KW   Lipid-binding; Lipid transport; Transport; Steroidogenesis;
KW   Mitochondrion; Transit peptide.
FT   TRANSIT       1      ?       MITOCHONDRION (POTENTIAL).
FT   CHAIN         ?    284       STEROIDOGENIC ACUTE REGULATORY PROTEIN.
FT   DOMAIN      170    174       POLY-ALA.
SQ   SEQUENCE   284 AA;  31626 MW;  ?80?15?1419B72 CRC64;


  Query Match             100.0%;  Score 1429;  DB 1;  Length 284;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-107;
  Matches  276;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
[illegible]   1 MEIATEK [illegible] SYE [illegible] MENME [illegible] EHLAV [illegible] NWEA [illegible]
[illegible]   1 MEIATEK [illegible] SYE [illegible] MENME [illegible] LEHLAV [illegible] NWEA [illegible]

[illegible]  61 [illegible] EATLY [illegible] SYI [illegible] VAMQKA [illegible] NNL [illegible] WE [illegible]
[illegible]  61 [illegible] EATLY [illegible] SYI [illegible] VAMQKA [illegible] NNL [illegible] WE [illegible]

[illegible] 121 [illegible] EVVVI [illegible] YEE [illegible] MEAM [illegible] WN [illegible] NVK [illegible]
[illegible] 121 LEVVVI [illegible] YEE [illegible] MEAM [illegible] WN [illegible] NVK [illegible]

[illegible] 181 KEEVSVE [illegible] KE [illegible] SI [illegible] VLA [illegible] MATHE [illegible] EMPE [illegible] V [illegible] MV [illegible]
[illegible] 181 KEEVSVR [illegible] KE [illegible] SI [illegible] VLA [illegible] MATHE [illegible] EMPE [illegible] V [illegible] MV [illegible]

[illegible] 241 [illegible] DLK [illegible] WE [illegible] VI [illegible] ANM [illegible] A [illegible]
[illegible] 241 [illegible] DLK [illegible] WE [illegible] VI [illegible] AN [illegible] A [illegible]

RESULT 2
STAR_RAT
ID   STAR_RAT     STANDARD;     [illegible]     [illegible] AA.
AC   P97826; [illegible]
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN [illegible]
GN   STAR.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; [illegible]
OC   Mammalia; Eutheria; [illegible]
OX   NCBI_TaxID=[illegible]
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=WISTAR; [illegible]
RX   MEDLINE=[illegible]
RA   Mizutani T., [illegible]
RT   "Molecular [illegible]
RT   steroidogenic acute regulatory [illegible] (StAR) [illegible]"
RL   [illegible] 61:[illegible]
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY; [illegible]
RX   MEDLINE=[illegible]; PubMed=[illegible]
RA   Lee H.K., Ahn [illegible] Kang [illegible]
RT   "Nucleotide sequence [illegible]
RT   complementary DNA.";
RL   Biochem. Biophys. Res. Commun. [illegible]
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY; [illegible]
RX   MEDLINE=[illegible]; PubMed=[illegible]
RA   Ariyoshi N., [illegible]
RA   [illegible] R.;
RT   "Characterization [illegible]
RT   [illegible] kb base pair mRNA [illegible]
RT   precedes [illegible] of [illegible] mRNA and protein.";
RL   J. Biol. Chem. [illegible]
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY; [illegible]
RX   MEDLINE=[illegible]
RA   Furukawa A., [illegible]
RT   "Steroidogenic acute [illegible]
RT   constitutively expressed in the adult rat central [illegible]
RT   colocalization of StAR [illegible]
RT   hydroxysteroid dehydrogenase in the rat brain.";
RL   J. Neurochem. [illegible]
CC   -!- FUNCTION: PLAYS A KEY ROLE IN [illegible]
CC       ENHANCING THE METABOLISM OF [illegible]
CC       RANSFERRED [illegible]
```

```
CC       [illegible] MEMBRANE [illegible] MITOCHONDRIA [illegible]
CC   [illegible] MITOCHONDRIA.
CC   [illegible] FAMILY [illegible] STAR FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. [illegible]
CC   between the Swiss Institute of Bioinformatics and the EMBL [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible] agreement [illegible]
CC   [illegible]
CC   --------------------------------------------------------------------------
DR   EMBL; AB[illegible]
DR   EMBL; [illegible]
DR   EMBL; AB[illegible]
DR   EMBL; AB[illegible]
DR   InterPro; [illegible]
DR   InterPro; [illegible]
DR   Pfam; [illegible]
DR   PRINTS; [illegible]
KW   Lipid [illegible] Transport; Steroidogenesis;
KW   Mitochondrion; [illegible]
FT   TRANSIT      [illegible]      MITOCHONDRION (POTENTIAL).
FT   CHAIN        [illegible]      STEROIDOGENIC ACUTE REGULATORY PROTEIN.
FT   CONFLICT     [illegible]      M [illegible] AV [illegible] VI [illegible] MI [illegible]
FT                                  [illegible] VVVN [illegible]
FT                                  [illegible] (IN [illegible])
SQ   SEQUENCE   [illegible] AA;   [illegible] MW;   [illegible] CRC64;

  [illegible]
  [illegible]
  [illegible]

Qy   [illegible] M [illegible] MENME [illegible] AV [illegible] WM [illegible]
Db   [illegible] M [illegible] MENME [illegible] HLAV [illegible] WM [illegible]

Qy   [illegible] VAMQKA [illegible] NNL [illegible] WE [illegible] EV [illegible] KVN [illegible] V [illegible]
Db   [illegible] EAMQKA [illegible] NNL [illegible] WE [illegible] EV [illegible] KVN [illegible] V [illegible]

Qy   [illegible] MEAM [illegible] WN [illegible] NVK [illegible] AAA [illegible]
Db   [illegible] MEAM [illegible] WN [illegible] NVK [illegible] AAAAGN [illegible]

Qy   [illegible] VLA [illegible] MATHE [illegible] EMPE [illegible] MV [illegible] A [illegible] K [illegible] W [illegible]
Db   [illegible] VLA [illegible] MA [illegible] EMPE [illegible] MV [illegible] A [illegible] K [illegible] W [illegible]

Qy   [illegible] W [illegible] N [illegible] V [illegible] ANE [illegible]
Db   [illegible] W [illegible] N [illegible] V [illegible] A [illegible]

RESULT [illegible]
STAR_MESA[illegible]
ID   STAR_MESA[illegible]     STANDARD;     [illegible]     [illegible] AA.
AC   [illegible]
DT   [illegible] NOV [illegible] (Created)
DT   [illegible] NOV [illegible] (Last sequence update)
DT   [illegible] MAY [illegible] (Last annotation update)
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN [illegible]
GN   STAR.
OS   Mesocricetus [illegible]
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; [illegible] Muridae; [illegible]
OC   M[illegible]
OX   NCBI_TaxID=[illegible]
RN   [illegible]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=[illegible]
RA   [illegible] A. [illegible] M. [illegible] A. [illegible]
RA   Lehoux J.-G.;
RT   "Adrenocorticotropin regulates the lev[illegible] of the steroidogenic acute
RT   regulatory (StAR) protein mRNA in hamster adrenals.";
RL   Endocr. Res. 22:515-520(1996).
CC   -!- FUNCTION: PLAYS A KEY ROLE IN [illegible] SYNTHESIS BY
CC       ENHANCING THE METABOLISM OF CHOLESTEROL INTO PREGNENOLONE.
CC       TRANSPORTER THAT BINDS TO AND TRANS[illegible] CHOLESTEROL [illegible] THE
CC       INTERMEMBRANE SPACE OF THE MITOCH[illegible] (PROBABLE).
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL.
CC   -!- SIMILARITY: BELONGS TO THE STAR FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as [illegible] as its content is [illegible]
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement [illegible]
CC   or send an email to [illegible]
CC   --------------------------------------------------------------------------
DR   EMBL; U66490; AAB06763.1; -.
DR   InterPro; IPR000799; -.
DR   InterPro; IPR002913; -.
     Pfam; PF01852; START; 1.
     PRINTS; PR00978; STARPROTEIN.
     Lipid-binding; Lipid transport; Transport; Steroidogenesis;
KW   Mitochondrion; Transit peptide.
FT   TRANSIT       1      ?       MITOCHONDRION (POTENTIAL).
FT   CHAIN         ?    284       STEROIDOGENIC ACUTE REGULATORY PROTEIN.
SQ   SEQUENCE   284 AA;  31497 MW;  [illegible]

  Query Match             92.7%;  Score 1[illegible];  DB 1;  Length 284;
  Best Local Similarity   92.1%;  Pred. No. [illegible]e-98;
  Matches  255;  Conservative   11;  Mismatches   10;  Indels   [illegible];  Gaps   [illegible];

Qy          1 MFLATFKLCAGSSYRHMRNMKGLRHQAV[illegible]AI[illegible]NWRALCDSSDCW[illegible]
              [illegible]
Db          1 MLLATFKLCAGSSYRHLRNMKGLRHQAV[illegible]AI[illegible]NR[illegible]
 
Qy         61 QLEATLYSDQELSYIQQGEVAMQKALGI[illegible]NN[illegible]
              [illegible]
Db         61 QLEAALYS[illegible]QELSYIQQ[illegible]VAMQKALS[illegible]SN[illegible]

Qy        121 LEVVVD[illegible]PMERLYEELVERMEAM[illegible]EWNPNVKE[illegible]
              [illegible]
Db        121 LEVVVLQPMERLYAELVERMEAMGEWNPNVKE[illegible]

Qy        181 RDFVSVRCTKRRGSTCVLAGMATDFGNM[illegible]
              [illegible]
Db        181 RDFVSVRCAKRRGSTCVLAGIATHFGEMPEQ[illegible]

          241 LLSIDLKGWLPKTIINQV[illegible]
              [illegible]
Db        241 LLSID[illegible]KGWLPKSIINQVLSQTQ[illegible]MFFANHL[illegible]

RESULT     4
STAR_HORSE
ID   STAR_HORSE     STANDARD;      PRT;   2[illegible] AA.
AC   O46689;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN PRECURSOR.
GN   STAR.
OS   Equus caballus (Horse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX   NCBI_TaxID=9796;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Follicle;
RX   MEDLINE=99124482; PubMed=9927292
```

```
RA   Kerban A., Boerboom D., Sirois J.;
RT   "Human chorionic gonadotropin induces an inverse regulation of
RT   steroidogenic acute regulatory prot[illegible] messenger ribonucleic acid in
RT   theca interna and granulosa cells of [illegible] preovulatory follicles.";
RL   Endocrinology 140:667-674(1999).
CC   -!- FUNCTION: PLAYS A KEY ROLE IN [illegible] HORMONE SYNTHESIS BY
CC       ENHANCING THE METABOLISM OF CH[illegible] INTO [illegible]
CC       TRANSPORTER THAT BINDS TO AND TRANS[illegible] THE
CC       INTERMEMBRANE SPACE OF THE MIT[illegible] (PROBABLE).
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL.
CC   -!- SIMILARITY: BELONGS TO THE STAR FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. [illegible] through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement [illegible]
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AF031697; AAC04704.1; -.
DR   EMBL; AF031696; AAC04703.1; -.
DR   InterPro; IPR000799; -.
DR   InterPro; IPR002913; -.
DR   Pfam; PF01852; START; 1.
DR   PRINTS; PR00978; STARPROTEIN.
KW   Lipid binding; Lipid transport; Transport; Steroidogenesis;
KW   Mitochondrion; Transit peptide.
FT   TRANSIT       1      ?       MITOCHONDRION (POTENTIAL).
FT   CHAIN         ?    285       STEROIDOGENIC ACUTE REGULATORY PROTEIN
SQ   SEQUENCE   285 AA;  31852 MW;  [illegible]

  Query Match             88.3%;  Score [illegible];  DB 1;  Length 285;
  Best Local Similarity   87.7%;  Pred. No. 1.3e-93;
  Matches  24[illegible];  Conservative   16;  Mismatches   [illegible];  Indels   [illegible];  Gaps   1;

Qy          1 MFLATFK[illegible]YRHMRNMKGLRHQAV[illegible]WM[illegible]VRRRSS[illegible]LG 5[illegible]
              [illegible]
Db          1 MLLATFK[illegible]YRHVRNMKGLRHQAV[illegible]QVRRQSSLLG 6[illegible]

Qy         60 SQLEAT[illegible]VAMQKAL[illegible]EMV[illegible]VGKVF 119
              [illegible]
Db         61 SQLE[illegible]AMQKAL[illegible]VGKVF 120

Qy        120 RLEVVV[illegible]LYEELVERMLAMGEWN[illegible]AAAAGNLVG 179
              [illegible]
Db        121 RLEVLV[illegible]LYEELVERMLAMGEWN[illegible]AAA[illegible] 180

Qy        180 PRDFVSVR[illegible]KRRGSTCVLAGMATDFG[illegible]KLT 239
              [illegible]
Db        181 PRDFVSVR[illegible]ARRRGS[illegible]VLAGMA[illegible] 240

Qy        240 WLLSI[illegible]KGWLPK[illegible]INQVL[illegible] 276
              [illegible]
Db        241 WLLSIDLKGWLPKTIINQVLSQTQVDFANHLRKRLES 277

RESULT     5
STAR_PIG
ID   STAR_PIG       STANDARD;      PRT;   285 AA.
AC   Q28996; Q95[illegible];
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN PRECURSOR.
GN   STAR.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
```

```
RC   TISSUE=Ovary;
RX   MEDLINE=97[illegible]; PubMed=[illegible];
RA   [illegible] N., [illegible]
RA   Silverstrides [illegible]
RT   "[illegible]
RT   gene expression [illegible]"
RL   Endocrinology [illegible]
RN   [2]
RP   SEQUENCE OF [illegible] mRNA
RC   TISSUE=[illegible]
RX   MEDLINE=971[illegible]; PubMed=[illegible]
RA   Balasubramanian [illegible]
RA   Veldhuis J.D.
RT   "Regulation [illegible]
RT   protein (StAR) [illegible]
RT   follicle-stimulating [illegible]
RL   Endocrinology [illegible]
CC   -!- FUNCTION: PLAYS A KEY ROLE IN [illegible]
CC       ENHANCING THE METABOLISM OF [illegible]
CC       TRANSPORTER THAT [illegible] AND TRANSFERS [illegible]
CC       INTERMEMBRANE SPACE OF THE MITOCHONDRION [illegible]
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIA
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN ADULT AND [illegible]
CC       TISSUES, INCLUDING ADULT [illegible] AND OVARY [illegible]
CC       GLANDS AS WELL AS STEROIDOGENIC [illegible]
CC       DEVELOPING FETAL [illegible]
CC       THE FETAL OVARY.
CC   -!- SIMILARITY: BELONGS TO THE STAR FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright [illegible]
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. [illegible]
CC   use by non-profit institutions as long as [illegible]
CC   modified and this statement [illegible]
CC   entities requires a license agreement (See http://www.[illegible]
CC   or send an email to [illegible]
CC   --------------------------------------------------------------------------
DR   EMBL; [illegible]; AAC[illegible]
DR   EMBL; [illegible]; AAC[illegible]
DR   InterPro; IPR[illegible];
DR   InterPro; IPR00[illegible];
DR   Pfam; PF[illegible]; START;
DR   PRINTS; PR0[illegible]; STARPROTEIN.
KW   Lipid transport; [illegible]
KW   Mitochondrion; Transit peptide.
FT   TRANSIT       [illegible]       MITOCHONDRION [illegible]
FT   CHAIN         [illegible]  28[illegible]  STEROIDOGENIC ACUTE [illegible]
FT   CONFLICT      [illegible]
SQ   SEQUENCE   28[illegible] AA;  [illegible] MW;  [illegible]

  Query Match             [illegible]
  Best Local Similarity   [illegible]  Pred. No. [illegible]
  Matches  24[illegible];  Conservative  [illegible];  Mismatches  [illegible]

Qy          1 MLLATFKL[illegible]
              |||||||[illegible]
Db          1 MLLATFKL[illegible]

Qy         61 SGLEATIY[illegible]
              [illegible]
Db         61 SGLEDIEY[illegible]

Qy        120 RLEVVVDQLMEP[illegible]
              [illegible]
Db        121 RLEVVVDQLMEP[illegible]

Qy        180 PRGFVSVR[illegible]
              [illegible]
Db        181 PRGFVSVR[illegible]

Qy        240 WLLVQLKGWL[illegible]
              [illegible]
```

```
Db        [illegible]

RESULT [illegible]
STAR_HUMAN
ID   STAR_HUMAN     STANDARD;      PRT;   [illegible] AA.
AC   P49675; [illegible]
DT   [illegible] (Rel. [illegible], Created)
DT   [illegible] (Rel. [illegible], Last sequence update)
DT   [illegible] (Rel. [illegible], Last annotation update)
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN, MITOCHONDRIAL [illegible]
GN   STAR.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; [illegible]
OC   Mammalia; Eutheria; Primates; Catarrhini; [illegible]
OX   NCBI_TaxID=[illegible]
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Adrenal [illegible]
RX   MEDLINE=[illegible]; PubMed=[illegible]
RA   Sugawara T., [illegible]
RA   [illegible]
RA   [illegible]
RT   "[illegible]
RT   [illegible]
RT   [illegible]"
RL   Proc. Natl. Acad. [illegible]
RN   [2]
RP   SEQUENCE FROM N.A.
RC   [illegible]
RX   MEDLINE=[illegible]; PubMed=[illegible]
RA   Sugawara [illegible] Miller W.L.
RA   [illegible]
RT   "[illegible] protein (StAR)
RT   [illegible]
RT   [illegible]"
RL   [illegible]
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=[illegible]; PubMed=[illegible]
RA   [illegible]
RA   [illegible]
RT   "[illegible]
RT   [illegible]"
RL   [illegible]
RN   [4]
RP   SEQUENCE FROM N.A.
RC   [illegible]
RA   [illegible]
RL   [illegible] to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   VARIANTS [illegible]
RX   MEDLINE=[illegible]; PubMed=[illegible]
RA   [illegible] Sugawara T., [illegible] Miller W.L.
RT   "[illegible] adrenal
RT   hyperplasia";
RL   New Engl. J. Med. [illegible]
RN   [6]
RP   VARIANT [illegible]
RX   MEDLINE=[illegible]; PubMed=[illegible]
RA   [illegible]
RA   [illegible]
RT   "A [illegible]
RT   [illegible] regulatory protein gene in a Japanese patient
RT   with [illegible] hyperplasia."
RL   Hum. Mutat. Suppl. [illegible] (1998).
CC   -!- FUNCTION: PLAYS A KEY ROLE IN [illegible] SYNTHESIS BY
CC       ENHANCING THE METABOLISM OF CHOLESTEROL INTO PREGNENOLONE.
CC       [illegible] AND TRANSFER [illegible]
CC       INTERMEMBRANE SPACE OF THE MITOCHONDRION (PROBABLE).
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIA.
CC   -!- [illegible] EXPRESSED IN [illegible] AND [illegible]
CC       [illegible]
CC   -!- DISEASE: DEFECTS IN STAR CAUSE CONGENITAL LIPOID ADRENAL
         HYPERPLASIA (CLAH), THE MOST SEVERE FORM OF ADRENAL HYPERPLASIA.
CC   -!- SIMILARITY: BELONGS TO THE STAR FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement [illegible]
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; U17280; AAC50141.1; -.
DR   EMBL; U29105; AAC50234.1; -.
DR   EMBL; U29099; AAC50234.1; JOINED.
DR   EMBL; U29100; AAC50234.1; JOINED.
DR   EMBL; U29101; AAC50234.1; JOINED.
DR   EMBL; U29102; AAC50234.1; JOINED.
DR   EMBL; U29103; AAC50234.1; JOINED.
DR   EMBL; U29104; AAC50234.1; JOINED.
DR   EMBL; S79669; AAB35726.1; -.
DR   EMBL; AF035277; AAB88174.1; -.
     MIM; 600617; -.
     MIM; 201710; -.
     InterPro; IPR000799; -.
DR   InterPro; IPR002913; -.
DR   Pfam; PF01852; START; 1.
DR   PRINTS; PR00978; STARPROTEIN.
KW   Lipid-binding; Lipid transport; Transport; Steroidogenesis;
KW   Mitochondrion; Transit peptide; Disease mutation; Polymorphism.
FT   TRANSIT       1      ?       MITOCHONDRION (POTENTIAL).
FT   CHAIN         ?    285       STEROIDOGENIC ACUTE REGULATORY PROTEIN.
FT   VARIANT     182    182       R -> L (IN CLAH).
FT                                /FTId=VAR_05627.
FT   VARIANT     203    203       D -> A.
FT                                /FTId=VAR_05628.
SQ   SEQUENCE   285 AA;  31958 MW;  D3330[illegible]568[illegible]BAD CRC64;

  Query Match             87.6%;  Score [illegible];  DB [illegible];  Length [illegible];
  Best Local Similarity   87.0%;  Pred. No. [illegible];
  Matches  241;  Conservative   18;  Mismatches   17;  Indels    1;  Gaps    1;

Qy          1 MFLATFKLCAGSSYRHMRNMKGLRHQAVLA[illegible]NWRALGD[illegible] WMGQVRRRSSLLG 59
              | ||||||||||||||||||||||| |||:|| |      | |||    :|    :   |  |  | ||
Db          1 MLLATFKLCAGSSYRHMRNMKGLRQQAVMAI[illegible]NRRALGGPTPSTWINQVRRRSSLL[illegible]

Qy         60 SQLEATLYSDQELSYIQQGEVAMQKAL[illegible]
              |:|| ||||||||:|:||||  ||||||||       || |||||:|||: :||:||||||
Db         61 SRLEETLYSDQELAYLQQGEEAMQKAL[illegible]

Qy        120 RLEVVVDQPMDRLYEELVDRMEAMGEWNPNV[illegible]KVLQRIGKDTVITHELAAAAAGNLVG 179
              ||||||||||:||||||:|||||||||||| |  ||:|||| |:|| | | |:|
Db        121 RLEVVVDQPMERLYEELVERMEAMGEWNPNV[illegible]KV[illegible]QKIGKD[illegible]THELAA[illegible]AAGNLVG 180

Qy        180 PRDFVSVRCTKRRGSTCVLAGMATHFGEM[illegible]VIRAEHGPTCMVLHPLAGSPSKTKLT 239
              ||||||||| |||||||||||| | || |||    | ||||||||||| ||||||||||||
Db        181 PRDFVSVRCAKRRGSTCVLAGM[illegible]TD[illegible]GNMP[illegible]RAEHGPTCMVLHPLAGSPSKTKLT 240

Qy        240 WLLSIDLKGWLPKTIINQVLSQTQIEFANHLRKRLEA 276
              ||||||||||||:|||||||||||::| |||| |   :.
Db        241 WLLSIDLKGWLPKSIINQVLSQTQVDFANHLRKRLES 277


RESULT    7
STAR_BOVIN
ID   STAR_BOVIN     STANDARD;      PRT;   285 AA.
AC   Q28918;
DT   01-NOV-1997 (Rel. 35, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN PRECURSOR.
GN   STAR.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID 9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Corpus luteum;
RX   MEDLINE 96011827; PubMed 7488004;
RA   Hartung S., Rust W., Balvers M., Ivell R.;
RT   "Molecular cloning and in vivo expression of the bovine steroidogenic
RT   acute regulatory protein.";
RL   Biochem. Biophys. Res. Commun. 215:646-653(1995).
RN   [2]
RP   REVISIONS TO 85 AND 174.
RA   Hartung S., Rust W., Balvers M., Ivell R.;
RL   Submitted (MAY-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: PLAYS A KEY ROLE IN STEROID HORMONE SYNTHESIS BY
CC       ENHANCING THE METABOLISM OF CHOLESTEROL INTO PREGNENOLONE.
CC       TRANSPORTER THAT BINDS TO AND TRANSFERS CHOLESTEROL THROUGH THE
CC       INTERMEMBRANE SPACE OF THE MITOCHONDRION (PROBABLE).
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL.
CC   -!- TISSUE SPECIFICITY: CORPUS LUTEUM AND ADRENAL GLAND.
CC   -!- SIMILARITY: BELONGS TO THE STAR FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement [illegible]
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; Y17259; CAA76717.1; -.
DR   EMBL; Y17260; CAA76718.1; -.
DR   InterPro; IPR000799; -.
DR   InterPro; IPR002913; -.
DR   Pfam; PF01852; START; 1.
DR   PRINTS; PR00978; STARPROTEIN.
KW   Lipid-binding; Lipid transport; Transport; Steroidogenesis;
KW   Mitochondrion; Transit peptide.
FT   TRANSIT       1      ?       MITOCHONDRION (POTENTIAL).
FT   CHAIN         ?    285       STEROIDOGENIC ACUTE REGULATORY PROTEIN.
FT   CONFLICT     75     75       H -> ? (IN [illegible] SEQUENCE).
FT   CONFLICT     99     99       E -> K (IN GENOMIC SEQUENCE).
SQ   SEQUENCE   285 AA;  31845 MW;  [illegible] CRC64;

  Query Match             85.0%;  Score [illegible];  DB [illegible];  Length 285;
  Best Local Similarity   [illegible];  Pred. No. [illegible];
  Matches  235;  Conservative   21;  Mismatches   20;  Indels    1;  Gaps    1;

Qy          1 MFLATFKL[illegible]YRHMRNMKGLRHQAV[illegible]WMGQVRRRSSLLG 59
              | |||||||| |||||:|:||||: ||||     |  : |       : |   ||  ||
Db          1 MLLATFKL[illegible]SSYRHVRSMKGLQQQAV[illegible]AAWINQVRRRSSLL[illegible] 60

Qy         60 SQLEA[illegible]VAMQKAL[illegible]EVLSKMVPDVGKVF 119
              ||||  | |||||::||||| |||:||||  :|||  |||:   |||||||::||||||
Db         61 SQLEDPLYS[illegible]LLAHIQQGEEAMQRAL[illegible]WKKESRQ[illegible]DEVLSKV[illegible]PDV[illegible]KVF 120

Qy        120 RLEVVVDQPM[illegible]RLYEELVDRMEAMGEWN[illegible]KVLQRIGKD[illegible]ITHELAAAAA[illegible]NLVG 179
              ||||||:  :: |||||||:||||||  | |||    :|        |:||     : |:
Db        121 RLEVVVDQ[illegible]MERLYEELVERMEAMGEWN[illegible]KVL[illegible]ITHELAA[illegible]NLVG 180

Qy        180 PRDFVSVR[illegible]KRRGSTCVLAGMA[illegible]MVLHPLAGSPSKTKLT 239
              ||||||||:||||| ||||||||  :  |  |  |   | |         |   |  ||:|||
Db        181 PRDFVSVR[illegible]KRRG[illegible]VLAGMA[illegible]PLAGSPSKTKLT 240

Qy        240 WLLS[illegible]NQVLSQ[illegible] 276
              ||||:| |||  ||||||||||||||||::|    |  |:::
Db        241 WLLSIDLKGW[illegible]PKTIINQVLSQTQV[illegible] 277
```

```
RESULT [illegible]
STAR SHEEP
ID   STAR SHEEP    [illegible] AN [illegible] AA.
AC   P79245;
DT   30-MAY-2000 [illegible] Created)
DT   30-MAY-2000 [illegible]
DT   30-MAY-2000 [illegible]
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN (FRAGMENT).
GN   STAR.
OS   Ovis aries (Sheep).
OC   Eukaryota; Metazoa; [illegible]
OC   Mammalia; Eutheria; [illegible]
OC   Bovidae; Caprinae; [illegible]
OX   NCBI_TaxID=[illegible]
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Corpus [illegible]
RX   MEDLINE=[illegible] PubMed=[illegible]
RA   Juengel J.L., Meberg B.M., [illegible] A.M., [illegible]
RT   "Hormonal regulation of messenger ribonucleic [illegible]
RT   steroidogenic acute regulatory protein [illegible]"
RL   Endocrinology [illegible] 5420(1995).
CC   -!- FUNCTION: PLAYS A KEY ROLE IN [illegible]
CC       ENHANCING THE METABOLISM OF CHOLESTEROL INTO PREGNENOLONE
CC       TRANSPORTER THAT BINDS [illegible] AND TRANS[illegible]
CC       INTERMEMBRANE SPACE OF THE MITOCHONDRION [illegible]
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL.
CC   -!- SIMILARITY: BELONGS TO THE STAR FAMILY.
CC
CC   This SWISS-PROT entry [illegible]
CC   between the Swiss [illegible]
CC   the European Bioinformatics [illegible]
CC   use by non-profit institutions [illegible]
CC   modified and this statement [illegible]
CC   entities requires a license agreement [illegible]
CC   or send an email to [illegible]
CC
DR   EMBL; [illegible] AAB1705[illegible]
DR   InterPro; IPR0[illegible]
DR   PRINTS; PR00978; STARPROTEIN.
KW   Lipid-binding; Lipid transport; Transport; [illegible]
KW   Mitochondrion.
FT   NON_TER [illegible]
SQ   SEQUENCE   121 AA;   [illegible] MW;   [illegible]

  Query Match    [illegible]
  Best Local Similarity [illegible]
  Matches [illegible]  Conservative [illegible]  Mismatches [illegible]  Gaps [illegible]

Qy         96 KKEQQENCGEV[illegible]KM[illegible]VGKVERLLVVVL[illegible]
                 |      |
Db          1 KKENRQANCGEV[illegible]EV[illegible]VGKVERLLVVVL[illegible]

Qy        156 LQKIGKDIVI[illegible]LAA[illegible]AGNLVQP[illegible]
                 |   ||     |
Db         61 LQKIGKDI[illegible]AA[illegible]AGN[illegible]

Qy        216 R[illegible]
              |
Db        121 R[illegible]


RESULT 9
MLN64 HUMAN
ID   MLN64 HUMAN    STANDARD;    PRT;   [illegible] AA.
AC   Q14849;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   [illegible] (Rel. [illegible], Last annotation update)
DE   MLN 64 PROTEIN (CAB1 PROTEIN).
GN   MLN64 OR CAB1.
OS   Homo sapiens (Human).
```

```
OC   [illegible]
OC   Mammalia; [illegible]
OX   [illegible]
RN   [illegible]
RP   SEQUENCE FROM N.A.
RC   [illegible]
RX   [illegible]
RA   [illegible]
RA   [illegible]
RT   [illegible]
RT   [illegible]
RT   [illegible]
RL   [illegible]
RN   [illegible]
RP   SEQUENCE FROM N.A.
RC   [illegible]
RX   [illegible]
RA   Akiyama N., Sasaki H., [illegible] Sakamoto H., [illegible]
RA   [illegible]
RT   "Isolation of a [illegible] (ARL[illegible]
RT   [illegible] cDNA [illegible]
RT   [illegible]"
RL   [illegible]
RN   [illegible]
RP   X-RAY [illegible] AND [illegible]
RX   MEDLINE=[illegible]
RA   [illegible]
RT   "[illegible] STAR-related [illegible]"
RL   Nat. [illegible] 7:43-[illegible]
CC   -!- FUNCTION: [illegible] TRANS[illegible]
CC       STEROIDOGENESIS IN PLACENTA AND BRAIN.
CC   -!- SIMILARITY: [illegible] THE STAR FAMILY.
CC
CC   This SWISS-PROT entry [illegible]
CC   between the Swiss [illegible] EMBL [illegible]
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use [illegible] way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement [illegible]
CC   or send an email to license@isb-sib.ch).
CC
DR   EMBL; X80[illegible] CAA[illegible]
DR   EMBL; [illegible] BAA[illegible]
DR   PIR; [illegible]
DR   InterPro; [illegible]
DR   InterPro; [illegible]
DR   Pfam; PF01852; START; 1.
DR   PRINTS; PR00978; STARPROTEIN.
KW   Lipid-binding; Lipid transport; Transport; Steroidogenesis;
KW   [illegible]
SQ   SEQUENCE   [illegible] AA;   [illegible] MW;   [illegible]

  Query Match    [illegible]
  Best Local Similarity [illegible]  Pred. No. [illegible]
  Matches [illegible]  Conservative [illegible]  Mismatches [illegible]  Indels [illegible]  Gaps [illegible]

Qy   [illegible] AMLKA[illegible]NNLQGWKK[illegible]VEGKVLVVGKVERLLVVV[illegible]

Db   [illegible] AVVLLLALTNWKEEKNN[illegible]VYIEIVIEHHGF[illegible]

Qy   [illegible] AM[illegible]WN[illegible]NVK[illegible]KV[illegible]LAAAAA[illegible]NLV[illegible]

Db   [illegible]

Qy   [illegible]

Db   [illegible]NG[illegible]

Qy   [illegible]

Db   [illegible]
```

```
RESULT  10
ML64_MOUSE
ID   ML64_MOUSE     STANDARD;      PRT;   [illegible] AA.
AC   Q61542;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   MLN 64 PROTEIN (ES 64 PROTEIN).
GN   MLN64 OR ES64.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID 10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96039245; PubMed 7490069;
RA   Tomasetto C., Regnier C.H., Moog-Lutz C., Mattei M.-G., [illegible]
RA   Lidereau R., Basset P., Rio M.-C.;
RT   "Identification of four novel human genes amplified and overexpressed
RT   in breast carcinoma and localized to the q11-q21.3 region of
RT   chromosome 17.";
     Genomics 28:367-376(1995).
     -!- FUNCTION: BINDS AND TRANSPORTS CHOLESTEROL. PROMOTES
         STEROIDOGENESIS IN PLACENTA AND BRAIN (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE STAR FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; X82457; CAA57834.1; -.
DR   InterPro; IPR000799; -.
DR   InterPro; IPR002913; -.
DR   Pfam; PF01852; START; 1.
DR   PRINTS; PR00978; STARPROTEIN.
KW   Lipid binding, Lipid transport, Transport, Steroidogenesis.
SQ   SEQUENCE   446 AA;  50469 MW;  DFF43[illegible]4F3E1E2 CRC64;

  Query Match             25.7%;  Score [illegible];  DB 1;  Length [illegible];
  Best Local Similarity   35.1%;  Pred. No. [illegible]e-22;
  Matches [illegible];  Conservative [illegible];  Mismatches [illegible];  Indels [illegible];  Gaps [illegible];

Qy         23 LRHQAVLAICQFLNWPALCDS---SPGWMQ[illegible]
Db        186 [illegible]

           73 SYIQQGEVAMQKAL[illegible]NNQ[illegible]WEFEE[illegible]
Db        238 EYIRQGKEATAVVDQILAQEENWKFEK[illegible]VY[illegible]VPFHCKIPILKIPL[illegible]PARLV 297

Qy        133 YEELVDPMEAMCEWNPNVKEIKVLQRI[illegible]HELAAAAAGNLV[illegible]
Db        298 YQEV[illegible]VIWNETVTACQ[illegible]

Qy        193 GS[illegible]VLACMATHF[illegible]EMPEQS[illegible]VIRAFHGP[illegible]
Db        358 [illegible]

Qy        253 T[illegible]NQVL[illegible]QIQ[illegible]EFANHLRKRL 274
Db        418 YLIHQSLGATMFEEFAFHLRQRV 439


RESULT  11
YX7O_CAEEL
ID   YX7O_CAEEL     STANDARD;      PRT;   [illegible] AA.
AC   O17883;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   HYPOTHETICAL 27.1 KDA PROTEIN F52F12.7 IN CHROMOSOME [illegible]
GN   F52F12.7.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID 6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   Matthews L.;
RL   Submitted ([illegible]-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: BELONGS TO THE STAR FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; Z81228; CAB03736.1; -.
DR   WormPep; F52F12.7; CE10852.
DR   InterPro; IPR000799; -.
DR   InterPro; IPR002913; -.
DR   Pfam; PF01852; START; 1.
DR   PRINTS; PR00978; STARPROTEIN.
KW   Hypothetical protein.
SQ   SEQUENCE   241 AA;  27124 MW;  [illegible] CRC64;

  Query Match             12.7%;  Score [illegible];  DB 1;  Length 241;
  Best Local Similarity   28.8%;  Pred. No. [illegible]e-07;
  Matches   65;  Conservative   33;  Mismatches   78;  Indels   5[illegible];  Gaps   12;

Qy         78 GEVAMQKALGILNNQE[illegible]GWKKE[illegible]
Db         33 GEV FAEVEATENDENYLSHAGWEKDESN[illegible]VTAKD[illegible]MVI[illegible]PMPVEIV 90

Qy        133 YEELVDPMEAMCEWNPNVKEIKVLQR[illegible]LAAAAAGN---LVGPREFV 184
Db         91 MKEIWN[illegible]AA[illegible]VVGPREFV 140

Qy        185 [illegible]VLACMA[illegible]
Db        141 EARIWRKV[illegible]LASRCVTVPT[illegible]VL[illegible]NPENPEI [illegible]2

Qy        236 TK[illegible]WL[illegible]GWLF[illegible]NQVL[illegible]
Db        193 [illegible]VVNL[illegible]KGYLEKMLVNQVIGE[illegible] 238


RESULT  12
FBX2_HUMAN
ID   FBX2_HUMAN     STANDARD;      PRT;   29[illegible] AA.
AC   Q9UK22, Q9[illegible];
DT   01-OCT-2000 (Rel. 40, Created)
DT   01-OCT-2000 (Rel. 40, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   F-BOX ONLY PROTEIN 2.
GN   FBXO2 OR FBX2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID 9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   PubMed 105[illegible];
RA   Winston J.T., Koepp D.M., Zhu [illegible], [illegible], Harper J.W.;
RT   "A family of mammalian F-box proteins.";
```

```
RL   [illegible]
RN   [2]
RP   SEQUENCE [illegible] M N A
RA   [illegible] M
RA   [illegible] M.;
RT   "Identification of [illegible]
RL   Curr. Biol. [illegible]
CC   -!- FUNCTION: [illegible]
CC       PROTEINS AND PROMOTES THE [illegible]
CC   -!- SUBUNIT: PART OF A [illegible]
CC       (BY SIMILARITY).
CC   -!- SIMILARITY: CONTAINS [illegible] DOMAIN.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. [illegible]
CC   between the Swiss Institute of [illegible]
CC   the European Bioinformatics Institute. [illegible]
CC   use by non-profit institutions [illegible]
CC   modified and this statement is not removed. [illegible]
CC   entities requires a license agreement [illegible]
CC   or send an email to [illegible]
CC   --------------------------------------------------------------------------
DR   EMBL; AF187[illegible]; AAF[illegible]
DR   EMBL; AF174[illegible]; AAF[illegible]
DR   InterPro; IPR00[illegible]
DR   Pfam; PF006[illegible]; [illegible]
DR   PROSITE; PS[illegible]
KW   Ubiquitin conjugation.
FT   DOMAIN [illegible]
FT   CONFLICT [illegible]
FT   CONFLICT [illegible]
SQ   SEQUENCE [illegible] AA; [illegible] MW; [illegible]

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches [illegible]; Conservative [illegible]; Mismatches [illegible]

Qy        [illegible] LNWRAI[illegible]WMG[illegible]
Db        [illegible] LEWKELV[illegible]
Qy        [illegible] ALG[illegible]LNNQ[illegible]KE[illegible]
Db        1[illegible] W[illegible]VEH[illegible]
Qy        1[illegible] EME [illegible] AM[illegible]NVE [illegible]KV[illegible]AAAA[illegible]
Db        1[illegible] L[illegible]PAIVVE W[illegible]
Qy        194 [illegible]VLAGMA[illegible]
Db        24[illegible] GWMEISH[illegible]
Qy        2[illegible] KL[illegible]N [illegible]
Db        28[illegible] AEVIN [illegible]

RESULT [illegible]3
RF3_HAEIN
ID   RF3_HAEIN      STANDARD; [illegible] AA.
AC   P43928;
DT   01-NOV-1995 (Rel. [illegible], Created)
DT   01-NOV-1995 (Rel. [illegible], Last sequence update)
DT   15-DEC-1998 (Rel. [illegible], Last annotation update)
DE   PEPTIDE CHAIN RELEASE [illegible]
GN   PRFC OR HI17[illegible]
OS   Haemophilus influenzae.
OC   Bacteria; Proteobacteria; gamma [illegible]
OC   Haemophilus.
OX   NCBI_TaxID=72[illegible];
RN   [1]
RP   SEQUENCE FROM N.A.
```

```
RP   [illegible]
RX   [illegible]
RA   [illegible]
RA   [illegible] M.
RA   [illegible]
RA   [illegible] M.
RA   [illegible] M.
RA   [illegible] M. [illegible]
RA   [illegible] M.
RA   [illegible] M. [illegible]
RA   Venter [illegible]
RT   "Whole-genome [illegible] Haemophilus
RT   [illegible]";
RL   [illegible]
CC   -!- FUNCTION: INCREASES THE FORMATION OF [illegible] TERMINATION
CC       [illegible] ANTICIPATES [illegible] AND [illegible]
CC       [illegible] AND HAS [illegible]
CC       IT MAY INTERACT [illegible] WITH THE [illegible]
CC       [illegible] AND [illegible] SIGNIFICANTLY REDUCED BY GTP AND GDP [illegible]
CC       [illegible] (BY SIMILARITY).
CC   -!- [illegible] LOCATION: [illegible]
CC   -!- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
CC       [illegible] SUBFAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. [illegible]
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation
CC   the European Bioinformatics Institute. [illegible]
CC   [illegible]
CC   modified and this statement is not removed. [illegible]
CC   [illegible]
CC   [illegible]
CC   --------------------------------------------------------------------------
DR   EMBL; [illegible]; AAC[illegible]
DR   [illegible]
DR   [illegible]
DR   [illegible]
DR   [illegible]
KW   [illegible]
FT   NP_BIND [illegible] (BY SIMILARITY).
FT   NP_BIND [illegible] (BY SIMILARITY).
FT   NP_BIND [illegible] (BY SIMILARITY).
SQ   SEQUENCE [illegible] AA; [illegible] MW; [illegible]

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches [illegible]; Mismatches [illegible]

Qy        [illegible]
Db        [illegible]
Qy        [illegible]
Db        [illegible]
Qy        [illegible]
Db        [illegible]
Qy        [illegible]
Db        [illegible]
Qy        [illegible]
Db        [illegible]

RESULT [illegible]
[illegible]
ID   [illegible]
AC   [illegible]
```

```
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   TERMINASE LARGE SUBUNIT (G2P).
GN   2.
OS   Bacteriophage SPP1.
OC   Viruses; dsDNA viruses, no RNA stage; Tailed phages; Siphoviridae;
OC   Lambda phage group.
OX   NCBI_TaxID=10724;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92194332; PubMed=1548711;
RA   Chai S., Bravo A., Lueder G., Nedlin A., Trautner T.A., Alonso J.C.;
RT   "Molecular analysis of the Bacillus subtilis bacteriophage SPP1
RT   region encompassing genes 1 to 6. The products of gene 1 and gene 2
RT   are required for pac cleavage.";
RL   J. Mol. Biol. 224:87-102(1992).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Alonso J.C., Luder G., Stiege A.C., Chai S., Weise F., Trautner T.A.;
RL   Submitted (MAY-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: ESSENTIAL FOR DNA MATURATION AND PACKAGING.
CC   -!- SUBUNIT: DIMER OF A SMALL AND A LARGE SUBUNIT (POTENTIAL).
CC   -!- SIMILARITY: TO B.SUBTILIS XTMB AND [illegible].
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as  long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement [illegible]
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; X56064; CAA39537.1; -.
DR   EMBL; X97918; CAA66573.1; -.
KW   DNA packaging
FT   MUTAGEN      18     18       W->R: IN [illegible]; TEMPERATURE SENSITIVE.
FT   MUTAGEN      73     73       F->L: IN SUS19.
FT   MUTAGEN     233    233       L->F: IN SUS19.
FT   MUTAGEN     412    412       D->N: IN [illegible]; TEMPERATURE SENSITIVE.
SQ   SEQUENCE   422 AA;  48841 MW;  [illegible] CRC64;

  Query Match              6.3%;  Score [illegible];  DB 1;  Length [illegible];
  Best Local Similarity   22.3%;  Pred. No. [illegible];
  Matches   57;  Conservative   37;  Mismatches   99;  Indels [illegible];  Gaps [illegible];

Qy         48 MGQVRRRSSLLCSQLEA  ILYSDQELSY[illegible]AMGK     AL[illegible]NNQE[illegible]WKKES[illegible]
              | :||            ||   |: : | |  : :          *     |: |:
Db          1 MKEVELSEEF[illegible]LVW[illegible]VKAAQ[illegible]V[illegible]AMW[illegible]                 [illegible]

Qy        102 ENGDEVLSEMV[illegible]VCKVP[illegible]LEVVV[illegible]QPM[illegible]V[illegible]SMEAMC[illegible]NVK[illegible]V[illegible]
                   :|  |:|    |  |:    |:|          :   | :|      |           :
Db         49 ----LLMMMMLITFLVIRRVYNIVEQSVPEQ[illegible]A[illegible]MLEVGHEW                  KV[illegible]

Qy        161 KDTVITHELAAAAAGNLV    GPKDFVSVR[illegible]VLAGMAT[illegible]GEM[illegible]            [illegible]
              |  :     |      || :     |   |    :             :|||       |:
Db         93 FSPL---R[illegible]TYI[illegible]NSI[illegible]PV[illegible]KTI    ASKTFVACM   WIEELAEPKIEERV 14[illegible]

Qy        211 ---QS[illegible]VIPAF[illegible]PTCMVLHPLAC[illegible]WL[illegible]            [illegible]
                 :  |:|||  | |   :     : :| | |  : :   :       :||              : |
Db        14[illegible] SVIP[illegible]VLEACLP[illegible]ANTFVP[illegible]

Qy        260 SQTQIEEANHLPKRLE 275
              |:   |   |  :::| |
Db        206 SKAFIEEAEEVKRRNE 221


RESULT 15
HYES_MOUSE
ID   HYES_MOUSE      STANDARD;      PRT;   [illegible] AA.
AC   P34914;
DT   01-FEB-1994 (Rel. 28, Created)
```

```
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   SOLUBLE EPOXIDE HYDROLASE (SEH) (EC [illegible]) (EPOXIDE HYDRATASE)
DE   (CYTOSOLIC EPOXIDE HYDROLASE) (CEH
GN   EPHX2 OR EPH2.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE
RC   TISSUE=Liver;
RX   MEDLINE=[illegible]; PubMed=[illegible]34964;
RA   Grant D.F., Storms D.H., Hammock B.D.;
RT   "Molecular cloning and expression of murine liver soluble epoxide
RT   hydrolase.";
RL   J. Biol. Chem. [illegible]:17628-17633(1[illegible]
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NMRI;
RX   MEDLINE=96[illegible]; PubMed=[illegible]750907
RA   Johansson C., Stark A., Sandberg M., Ek B., Rask L., Meijer J.;
RT   "Tissue specific basal expression of soluble murine epoxide hydrolase
RT   and effects of clofibrate on the mRNA levels in extrahepatic tissues
RT   and liver.";
RL   Arch. Toxicol. 70:61-63(1995).
RN   [3]
RP   X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS)
RC   TISSUE=Liver;
RX   MEDLINE=99415911; PubMed=10485878;
RA   Argiriadi M.A., Morisseau C., Hammock B.D., Christianson D.W.;
RT   "Detoxification of environmental mutagens and carcinogens: structure,
RT   mechanism, and evolution of liver epoxide hydrolase.";
RL   Proc. Natl. Acad. Sci. U.S.A. 96:10[illegible](1999).
CC   -!- FUNCTION: THIS ENZYME ACTS ON EPOXIDES (ALKENE OXIDES, OXIRANES)
CC       AND ARENE OXIDES. PLAYS A ROLE IN XENOBIOTIC METABOLISM
CC       BY DEGRADING POTENTIAL TOXIC EPOXIDES. ALSO DETERMINES STEADY
CC       STATE LEVELS OF PHYSIOLOGICAL MEDIATORS.
CC   -!- CATALYTIC ACTIVITY: AN EPOXIDE + H(2)O = A GLYCOL.
CC   -!- SUBUNIT: HOMODIMER.
CC   -!- SUBCELLULAR LOCATION: CYTOPLASM AND PEROXISOMAL.
CC   -!- INDUCTION: BY COMPOUNDS THAT CAUSE PEROXISOME PROLIFERATION
CC       SUCH AS CLOFIBRATE, TIADENOL AND FENOFIBRATE.
CC   -!- PTM: THE N-TERMINUS IS BLOCKED.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as  long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; L05781; AAA37555.1; -.
DR   EMBL; Z37107; CAA85471.1; -.
DR   PIR; A47504; A47504.
DR   PDB; 1CQZ; 19-NOV-99.
DR   PDB; 1CR6; 19-NOV-99.
DR   SWISS-2DPAGE; P34914; MOUSE.
DR   MGD; MGI:[illegible]; Ephx2.
DR   InterPro; IPR000073; -.
DR   InterPro; IPR000639; -.
DR   InterPro; IPR0[illegible]4; -.
DR   Pfam; PF00702; Hydrolase; 1.
DR   Pfam; PF00561; abhydrolase; 1.
DR   PRINTS; PR00412; EPOXHYDRLASE.
DR   PRINTS; PR00413; HADHALOGNASE.
DR   PROSITE; PS00342; MICROBODIES_CTER; 1.
KW   Hydrolase; Peroxisome; Detoxification; [illegible]
KW   Aromatic hydrocarbons catabolism; 3D-structure.
FT   ACT_SITE    333    333       BY SIMILARITY.
FT   ACT_SITE    495    495       BY SIMILARITY.
FT   ACT_SITE    523    523       BY SIMILARITY.
FT   SITE        [illegible]
FT   CONFLICT    [illegible]
SQ   SEQUENCE    [illegible] MW [illegible]

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches [illegible]

Qy   [illegible] WEAL[illegible]WM[illegible]QV[illegible]
                     |
Db   [illegible] WVE[illegible]MDE[illegible] YE[illegible]AN[illegible]

Qy   [illegible] AL[illegible]ILNN [illegible]WK[illegible]EV[illegible]EMV[illegible]
                      : | |       | | |
Db   [illegible] [illegible]IVINNWE[illegible]

Qy   [illegible] EV[illegible]MEAN[illegible]NV [illegible]
                 | : | : : |
Db   16[illegible] [illegible]KA [illegible]NEV[illegible]

Qy   [illegible] EV[illegible]V[illegible]
                      |
Db   22[illegible] [illegible]V[illegible]N[illegible]V [illegible]YV [illegible]

Qy   24[illegible] W [illegible]

Db   2[illegible] WEY[illegible]PA[illegible]A [illegible]EVI [illegible]MF[illegible]


Search completed: [illegible]
[illegible] time [illegible]
```

```
                    GenCore version 4.5
             Copyright (c) 1993 - 20[illegible]  Compugen Ltd.

OM protein - protein search, using sw model

Run on:          August 22, 2001, 14:24:06 ; Search time 64.49 Seconds
                                             (without alignments)
                                             32[illegible].007 Million cell updates/sec

Title:           US 08 659-254-2
Perfect score:   1429
Sequence:        1 MFLATFKLCAGSSYRHMRNM.............QVLSQIQIHEANSLKKLLA 27[illegible]

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters:      219241

[illegible]mum DB seq length: 0
[illegible]mum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       PIR_68:*
                 1:  pir1:*
                 2:  pir2:*
                 3:  pir3:*
                 4:  pir4:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                              SUMMARIES
                 %
Result           Query
  No.    Score   Match  Length DB  ID                          Description
---------------------------------------------------------------------------------
    1     1429   100.0     284  2  A55455                      steroidogenic acut
    2     1326    92.8     362  2  JC5386                      steroidogenic acut
    3   1252.5    87.6     285  2  I38243                      steroidogenic acut
    4   1214.5    85.0     285  2  JC431[illegible]                     steroidogenic acu
    5      374    26.2     445  2  I38027                      MLN 64 protein - h
    6      181    12.7     211  2  T2251[illegible]                     hypothetical prote
    7    134.5     9.4     478  2  T1617[illegible]                     hypothetical prote
    8    127.5     8.9     196  2  [illegible]                    [illegible]
    9      101     7.1     296  2  T2323[illegible]                     hypothetical prote
   10     97.5     6.8     243  2  H82581                      DNA repair protein
   11       97     6.8     268  2  T43486                      hypothetical prote
   12     96.5     6.8    1281  2  G83405                      hypothetical prote
   13       96     6.7     322  2  T25416                      hypothetical prote
   14       94     6.6     686  2  E96611                      probable [illegible]
   15     92.5     6.5   10223  2  T30225                      polyketide synthas
   16       92     6.4     346  2  C82156                      conserved hypothet
   17       92     6.4     483  2  G75392                      glycosyl hydrolase
   18       92     6.4     527  2  C64159                      translation [illegible]
   19     90.5     6.3    1469  2  H96622                      probable ABC trans
   20     89.5     6.3     422  2  S24451                      hypothetical prote
   21       89     6.2     268  2  G70865                      probable DNA [illegible]
   22       88     6.2     377  2  C85596                      hypothetical prote
   23       88     6.2     404  2  B453[illegible]                     putrescine transp
   24       88     6.2     679  2  T05041                      hypothetical prote
   25     87.5     6.1    1047  2  A59245                      HIRA protein - [illegible]
   26       87     6.1    1130  2  A69816                      DNA exonuclease [illegible]
   27     86.5     6.1     343  2  JC5436                      probable [illegible]
   28     86.5     6.1     636  2  JC4960                      DNA topoisomerase
   29     86.5     6.1     994  2  A32792                      [illegible]
   30     86.5     6.1    6260  2  T30228                      polyketide synthas
   31       86     6.0     394  2  S37051                      4-hydroxybenzoate
   32       86     6.0    1198  2  T42224                      ladder protein [illegible]
   33       86     6.0    1410  1  A57013                      early endosome ant
   34     85.5     6.0     469  2  B70428                      general secretion
   35     85.5     6.0     518  2  T05876                      cytochrome P450 ho
   36     85.5     6.0     546  2  T40258                      probable [illegible] complex
   37       85     5.9     450  2  G71[illegible]                     probable [illegible] prote
   38       85     5.9     554  2  A47[illegible]                     epoxide hydrolase
   39       85     5.9     751  2  D72338                      (p)ppGpp synthetas
   40       85     5.9     752  2  F84852                      hypothetical prote
   41       85     5.9    1122  2  T28140                      hypothetical prote
   42       85     5.9    3433  1  S28381                      dystrophin - human
   43     84.5     5.9     338  2  JC7161                      acyl-CoA hydrolase
   44     84.5     5.9     433  2  B71769                      hypothetical prote
   45     84.5     5.9    1108  2  E71104                      probable cell divi

                              ALIGNMENTS

RESULT    1
A55455
steroidogenic acute regulatory protein precursor, mitochondrial - mouse
C;Species: Mus musculus (house mouse)
C;Date: 10-Feb-1995 #sequence_revision 10-Feb-1995 #text_change 08-Oct-1999
C;Accession: A55455
R;Clark, B.J.; Wells, J.; King, S.R.; Stocco, D.M.
J. Biol. Chem. 269, 28314-28322, 1994
A;Title: The purification, cloning, and expression of a novel luteinizing hormone-ind
ry protein (StAR).
A;Reference number: A55455; MUID:950[illegible]
A;Accession: A55455
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-284 <CLA>
A;Cross-references: GB:L36062; GB:L36061 [illegible]; NID:g1236242; PIDN:AAB94783.1; P
C;Keywords: mitochondrion; phosphoprotein
F;1-47/Domain: transit peptide (mitochondrion) #status predicted <TRP>

  Query Match             100.0%;  Score 1429;  DB 2;  Length 284;
  Best Local Similarity   100.0%;  Pred. No. [illegible]e-107;
  Matches  276;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MFLATFKLCAGSSYRHMRNMKGLRHQAV[illegible]SWMQVPEESSLLGS 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MFLATFKLCAGSSYRHMRNMKGLRHQAV[illegible]SWMQVPEESSLLGS 60

Qy         61 QLEATLYS[illegible]CEVAMQKAL[illegible]VISKMVPDVGKVFR 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 QLEA[illegible]EVAMQKA[illegible]KMV[illegible]VFR 120

Qy        121 LEVVVDQPMDKLYEELVDRMEAMGEWNPNV[illegible]IPELAAAAAGNLVHR 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 LEVVVDQPMDKLYEELVDRMEAMGEWNPNVKEIKVLQKIGKDTVITHELAAAAAGNLVHR 180

Qy        181 RDFVSVR[illegible]VLAGMATH[illegible]SKIKCLW 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 RDFVSVRCAKRRGSTCVLAGMATHFGEMPEQSGVIRAEHGPTCMVLHPLAGSPSKTKLTW 240

Qy        241 LLSIDLKGWLPKTIINQVLSQTQIEFANHLRKRLEA 276
              ||||||||||||||||||||||||||||||||||||
Db        241 LLSIDLKGWLPKTIINQVLSQTQIEFANHLRKRLEA 276


RESULT    2
JC5386
steroidogenic acute regulatory protein - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 04-Jun-1997 #sequence_revision [illegible] #text_change [illegible] May 1999
C;Accession: JC5386
R;Lee, H.; Ahn, [illegible]; Kwon, H.; Soh, [illegible]
```

```
Biochem. Biophys. Res. [illegible]
A;Title: Nucleotide [illegible]
A;Reference number: [illegible] MUID [illegible]
A;Accession: [illegible]
A;Molecule type: mRNA
A;Residues: [illegible]
A;Cross-references: [illegible]
C;Comment: This [illegible]

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches [illegible] Conservative [illegible] Mismatches [illegible]

Qy          5 [illegible]CAG[illegible]YHMRN[illegible]QAVLA[illegible]NW[illegible]WM[illegible]
Db         83 [illegible]CAG[illegible]YHMRN[illegible]QAVLA[illegible]WM[illegible]

Qy         65 [illegible]YSDQEL[illegible]AMQKA[illegible]NN[illegible]WKK[illegible]
Db        143 [illegible]YSDQEL[illegible]AMQKA[illegible]NN[illegible]WKK[illegible]

Qy        125 VL[illegible]MD[illegible]YE[illegible]VDRM[illegible]AMG[illegible]WNPNVK[illegible]KV[illegible]
Db        203 [illegible]MD[illegible]YE[illegible]VDRM[illegible]AMGEWNPNVK[illegible]KV[illegible]

Qy        185 SV[illegible]IK[illegible]VLA[illegible]
Db        263 SV[illegible]IK[illegible]VLA[illegible]

Qy        245 [illegible]WLPK[illegible]VL[illegible]ANH[illegible]
Db        323 [illegible]KVWLPK[illegible]VL[illegible]


RESULT      3
I38248
steroidogenic acute regulatory protein [illegible]
C;Species: Homo sapiens (man)
C;Date: 29 May 1998 [illegible]
C;Accession: I38248[illegible]
R;[illegible]wara, [illegible]
Biochemistry 34, [illegible]
A;Title: Structure of the [illegible]
A;Reference number: [illegible]48[illegible]
A;Accession: I38248
A;Status: preliminary [illegible]
A;Molecule type: DNA
A;Residues: 1-285 [illegible]
A;Cross-references: EMBL:[illegible] NID[illegible]
R;Sugawara, [illegible]
Proc. Natl. Acad. Sci. U.S.A. [illegible]
A;Title: Human ster[illegible]
A;Reference number: [illegible] MUID[illegible]
A;Accession: [illegible]
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-285 [illegible]
A;Cross-references: [illegible]
C;Genetics:
A;Gene: STAR
A;Cross-references: [illegible]
A;Map position: 8p[illegible]
A;Introns: [illegible]

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches [illegible] Conservative [illegible]

Qy          1 MFLA[illegible]KI[illegible]HM[illegible]NM[illegible]
Db          1 MLLA[illegible]KI[illegible]HM[illegible]NM[illegible]
```

```
Qy        [illegible]VAM[illegible]NN[illegible]KMV[illegible]V[illegible]KV[illegible]
Db        [illegible]KVM[illegible]KVV[illegible]V[illegible]KV[illegible]

Qy        [illegible]WNPNVK[illegible]AAAAA[illegible]
Db        [illegible]WNPNVK[illegible]AA[illegible]AA[illegible]

Qy        [illegible]MV[illegible]
Db        [illegible]MV[illegible]

Qy        [illegible]ANH[illegible]
Db        [illegible]ANH[illegible]

RESULT      4
[illegible]

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches [illegible]

Qy        [illegible]
Db        [illegible]AAWIN[illegible]

Qy        [illegible]VAM[illegible]NN[illegible]KMV[illegible]
Db        [illegible]

Qy        [illegible]WNPNVK[illegible]AAAAA[illegible]
Db        [illegible]WNPNVK[illegible]

Qy        [illegible]MV[illegible]
Db        [illegible]MV[illegible]

Qy        [illegible]ANH[illegible]
Db        [illegible]ANH[illegible]

[illegible]
C;Species: [illegible]
C;Date: [illegible]
C;Accession: [illegible]
[illegible]
A;Title: [illegible]
A;Reference number: [illegible]
A;Accession: [illegible]
```

```
A;Status: preliminary; translated from [illegible]
A;Molecule type: mRNA
A;Residues: 1-445 <RES>
A;Cross-references: EMBL:X8[illegible]198; NID:g[illegible]; PIDN:CAA[illegible]; PID:g[illegible]
A;Note: submitted to the EMBL Data Library, [illegible]ly 1994
C;Genetics:
A;Gene: MLN64

  Query Match             26.2%;  Score 37[illegible];  DB [illegible];  Length 445;
  Best Local Similarity   38.5%;  Pred. No. [illegible]5e-22;
  Matches   80;  Conservative   42;  Mismatches   86;  Indels   [illegible];  Gaps   [illegible];

Qy       67 YS[illegible]VAM[illegible]FA[illegible] 1[illegible]
              :|  ||   ||:||: |          ||   :| ||  |     |  |  :  :|   ||  : |:   :
Db      231 FSAQERRYIRQGKEATAVVDQILAQEENWKFEKNN[illegible]YC[illegible]TVYI[illegible]EVP[illegible]HCK[illegible]LKI[illegible]P 29[illegible]

Qy      127 [illegible]PM[illegible]YE[illegible]LVDRMFAMGFWNPNVK[illegible] 1[illegible]
               |  :  :|:|::  :  | |   ||   |       ::|:::     |:|:::::| ||| :|  : |:|:|
Db      291 [illegible]PAELVYQEVILQPERMVLWNKTVTACQ[illegible]NTL[illegible]YDVGAGAAG[illegible]VVG[illegible]PVNV [illegible]

Qy      187 [illegible]CTKRRCST[illegible]VLACMAT[illegible]HFCEMPEQSGV[illegible]EAF[illegible]MV[illegible]HPLA[illegible] 2[illegible]
              |   :||     : :|:||      |        :|  :       ::    |  :         :|:  |
Db      351 RR[illegible]FRRPDRYLSS[illegible]IATSHSAKPPTHKYV[illegible]VLK[illegible]ASN[illegible]PV[illegible]WILN[illegible]DL 41[illegible]

Qy      247 K[illegible]W[illegible]P[illegible]NQVL[illegible]SQTQIEEANHLRKRL 274
              || ||:  ::| |:  |   |||  |||:|:
Db      411 KGRLPRYLIHQSLAATMFEEFAFHLRQRI 4[illegible]


RESULT     6
T22513
hypothetical protein F52F12.7 - Caenor[illegible]abditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15 Oct 1999 #sequence_revision 1[illegible] #text_change [illegible]
C;Accession: T22513
R;Matthews, L.
submitted to the EMBL Data Library, December 199[illegible]
A;Reference number: Z1957[illegible]
A;Accession: T22513
A;Status: preliminary; translated from [illegible]
A;Molecule type: DNA
A;Residues: 1-241 <WIL>
A;Cross-references: EMBL:Z83228; PIDN:CAB[illegible]6.1; GSPDB:GN00019; CESP:F52F12.7
A;Experimental source: clone F52F12
C;Genetics:
A;Gene: CESP:F52F12.7
A;Map position: 1
A;Introns: 24,[illegible]

  [illegible]ery Match             12.7%;  Score [illegible];  DB 2;  Length 241;
  [illegible]st Local Similarity   28.8%;  Pred. No. [illegible]e-07;
  Matches   65;  Conservative   33;  Mismatches   78;  Indels   [illegible];  Gaps   [illegible];

Qy       78 [illegible]FVAM[illegible]FA[illegible]NP[illegible]E       [illegible]WK[illegible]N [illegible] 1[illegible]
             |||     :    | |::        || |:          |   :|   |   |::   :     :   |:: :
Db       33 GFV-FAFVFAIFNDFNY[illegible]SHAGWFFDF[illegible]NNF[illegible]IPH [illegible]

Qy      133 YFFLVDRMFAMGFWNFNVKE[illegible]KVL[illegible]R[illegible]K[illegible]     V[illegible]HELAAAAA[illegible]N[illegible] 1[illegible]
              :|   :  |||: |||  |     |    ||         ::|:            ||      :     |:  |
Db       91 MKFTWNCMFALPEWN[illegible]N    INFAAR[illegible]AF[illegible]VTY        [illegible] 1[illegible]

Qy      185 SVRCTKRRCSICVLACMATHFGEM-PE[illegible]MVLHPLA[illegible]-----    -SPSK [illegible]
             | |   :: |     :||   :               :  |                  || | |          :
Db      141 SAF[illegible]WFFV[illegible]AC[illegible]SVTVF[illegible]FF[illegible]GF[illegible] 1[illegible]

Qy      236 TKLTWLLS[illegible]DLLKGW[illegible]KIIINQVL[illegible]QTQ[illegible]          [illegible]EANHLRKR 27[illegible]
             |      ::   ||||:|||  |:|||:  :                |  |
Db      193 TLTDVVMLADLKCY[illegible]KMIVNQV[illegible]CR[illegible]M[illegible] V [illegible]NLEAKR 2[illegible]


RESULT     7
```

```
T16170
hypothetical protein F26F4.4 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 20 Sep 1999 #sequence_revision [illegible] Sep [illegible] #text_change 2[illegible] Sep [illegible]
C;Accession: T16170
R;Fulton, L.
submitted to the EMBL Data Library, March 199[illegible]
A;Description: The sequence of C. elegans [illegible] F26F4
A;Reference number: Z18471
A;Accession: T1617[illegible]
A;Status: preliminary; translated from [illegible]
A;Molecule type: DNA
A;Residues: 1-478 <FUL>
A;Cross-references: EMBL:[illegible]; NID:[illegible]; PIDN:AAA[illegible]219.1; CESP:F26
A;Experimental source: strain Bristol N2
C;Genetics:
A;Gene: CESP:F26F4.4
A;Introns: [illegible]

  Query Match              9.4%;  Score [illegible];  DB [illegible];  Length 478;
  Best Local Similarity   21.6%;  Pred. No. [illegible];
  Matches   53;  Conservative   5[illegible];  Mismatches   98;  Indels   41;  Gaps   11;

Qy       53 RRSSL[illegible]LYS[illegible]QELSYIQQ[illegible]VAM[illegible]KAI[illegible]NQ[illegible]WKF[illegible]SQ[illegible]NG[illegible]VLSKMV 112
             ||      |:   ||          :|                    :   :  : ::   ::   : : :   : ::
Db      243 RRQQNSGN[illegible]SEAFTAVPSRVS              [illegible]VETA[illegible]YDEFF[illegible]AEESSDEEARSELL 292

Qy      113    PDVGKV[illegible]VVVDQPMDELYEELVDE[illegible] AMG[illegible]    - -NPNVK-    --   --FIK 154
               ||   ::|     :  :  :  ::::  :|  :  |     :|                :
Db      293 VPPDIKKLF  [illegible]TLKE[illegible]LDQV EELM [illegible]K[illegible]LRGANL[illegible]ENYILVIFGK 34[illegible]

Qy      1[illegible]5 VLQRI[illegible]KR[illegible]V[illegible]HELAAAAAGNLVGPP[illegible]MATHFGEM[illegible]QS 21[illegible]
             ::    :       |  :  :  ::|  |        :  ||  :                 |     |       ::
Db      34[illegible] MIAHL[illegible]AMR[illegible]NG 4[illegible]

Qy      213 G  VIRAFHS[illegible]MVLHPLAGK[illegible]SKIE[illegible]WL[illegible]RI [illegible]NQVLS[illegible]EEANHL 2[illegible]
                  ::|    :  |:   ::   |           |        :     :|:  :|:  ::    :          |:  |:
Db      4[illegible] NEKIVKG[illegible]N[illegible]MIK[illegible]LKDEA[illegible]WIE[illegible]K[illegible]VN    YESEHV 4[illegible]

Qy      271 LKRLE 27[illegible]
             ::  |
Db      464 KKMNE 4[illegible]


RESULT     8
T46357
hypothetical protein DKFZp434G2416.1 - [illegible]
C;Species: Homo sapiens (man)
C;Date: 04-Feb-2000 #sequence_revision [illegible] #text_change 04-Feb-2[illegible]
C;Accession: T46357
R;Koehrer, K.; Beyer, A.; Mewes, H.W.; [illegible]; Wiemann, S.
submitted to the Protein Sequence Database, January 2000
A;Reference number: Z23037
A;Accession: T46357
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-19[illegible] <AAA>
A;Cross-references: EMBL:AL137657
A;Experimental source: adult testis; clone DKFZp434G2416
C;Genetics:
A;Note: DKFZp434G2416.1

  Query Match              8.9%;  Score [illegible];  DB [illegible];  Length [illegible];
  Best Local Similarity   29.7%;  Pred. No. [illegible];
  Matches   30;  Conservative   19;  Mismatches   [illegible];  Indels   [illegible];  Gaps   [illegible];

Qy      177 LVG[illegible]VSV[illegible]TKRRCS[illegible]VLACMAT[illegible]M[illegible]GV[illegible]EAF[illegible]MVLHPLAGSL[illegible] 2[illegible]4
             |: |||   :    ||          ::  ||:           |  :  |  :                       |:    :
Db        3 LISPRD[illegible]VKKY-E[illegible]ISSNA[illegible]GVR[illegible]NH[illegible]EPLPGEE[illegible] 6[illegible]

Qy      2[illegible] K[illegible]WLL[illegible]LKGWLPKI[illegible]INQVL[illegible]    AN[illegible]KRLE [illegible]
```

```
Db       62 KINLVIE[illegible]NVV[illegible]

RESULT 9
T23439
hypothetical protein K02[illegible]
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision [illegible]
C;Accession: T23[illegible]
R;Bardill, S.
submitted to the EMBL Data Library, March [illegible]
A;Reference number: [illegible]
A;Accession: T23[illegible]
A;Status: preliminary; translated from [illegible]
A;Molecule type: DNA
A;Residues: 1-29[illegible]
A;Cross-references: EMBL:[illegible] PIDN:AA[illegible]
A;Experimental source: [illegible]
C;Genetics:
A;Gene: CESP:K02[illegible]
A;Map position: X
A;Introns: 49/3; [illegible]

  Query Match [illegible]
  Best Local Similarity [illegible] Pred. No. [illegible]
  Matches [illegible] Conservative [illegible] Mismatches [illegible]

Qy      145 EWNPNVKEIKV[illegible]VIIHE[illegible]AAAA[illegible]NEV[illegible]

Db      114 KW[illegible]QSA[illegible]VIIHK[illegible]MM[illegible]VVMV[illegible]

Qy      201 MA[illegible]HEGEM[illegible]VI[illegible]MV[illegible]A[illegible]IWI[illegible]

Db      203 [illegible]VE[illegible]EK[illegible]VVV[illegible]MIN[illegible]

Qy      25[illegible] [illegible]QIEEANE[illegible]K[illegible]

Db      258 MEE[illegible]IA[illegible]K[illegible]


RESULT 10
H82581
DNA repair protein Rec[illegible] XF[illegible] [imported] - Xylella fastidiosa [illegible]
C;Species: Xylella fastidiosa
C;Date: 18 Aug 2000 #sequence_revision 20 Aug [illegible]
C;Accession: H82581
R;anonymous, The Xylella fastidiosa Consortium of the [illegible]
Nature 406, 151-157, 2000
A;Title: The genome sequence of the plant pathogen Xylella [illegible]
A;Reference number: A82515; MUID:[illegible]
A;Note: for a complete list of authors [illegible]
A;Accession: H82581
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-243 [illegible]
A;Cross-references: GB:AE[illegible] GB:AE[illegible] NID:[illegible]
A;Experimental source: strain [illegible]
R;Simpson, A.J.G.; [illegible]
Briones, M.R.[illegible]; Camargo, A.A.[illegible]
da Neto, E.; [illegible]
submitted to GenBank[illegible]
A;Authors: Ferreira, V.[illegible]A.[illegible]
[illegible] Marques, M.[illegible]
Machado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; [illegible]
A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; [illegible]
[illegible] Nunes, L.R.; Oliveira, M.A.[illegible]
Rodrigues, V.; Rosa, A.[illegible]
A;Authors: da Silva, A.C.R.[illegible]
M.; Tsuhako, M.H.; Vettore, A.[illegible] Vinhas[illegible] M.A.[illegible]
A;Reference number: A8[illegible]
A;Contents: annotation[illegible]
```

```
[illegible]
A;[illegible]
[illegible]

  Query Match [illegible]
  Best Local Similarity [illegible] N[illegible]
  Matches [illegible] Mismatches [illegible]

Qy      [illegible]NMK[illegible]AV[illegible]NWRA[illegible]WM[illegible]VE[illegible]IA[illegible]

Db      [illegible]EK[illegible]AI[illegible]AI[illegible]AEKKIAMIA[illegible]

Qy      [illegible]AM[illegible]VVV[illegible]

Db      [illegible]N[illegible]AVNE[illegible]

Qy      [illegible]NV[illegible]

Db      [illegible]AW[illegible]AN[illegible]

Qy      [illegible]VIA[illegible]MAI[illegible]MV[illegible]

Db      [illegible]HNNA[illegible]AI[illegible]MEAVI[illegible]

Qy      [illegible]

Db      [illegible]


RESULT [illegible]
[illegible]
hypothetical protein [illegible]
C;Species: Homo sapiens (man)
C;Date: [illegible]
C;Accession: [illegible]
R;[illegible]
submitted to the [illegible]
A;Reference number: [illegible]
A;Accession: [illegible]
A;Status: [illegible]
A;Molecule type: mRNA
A;Residues: [illegible]
A;Cross-references: EMBL:AL[illegible]
A;Experimental source: [illegible]
C;Genetics:
A;Note: [illegible]

  Query Match [illegible]
  Best Local Similarity [illegible] Pred. No. [illegible]
  Matches [illegible] Mismatches [illegible] Gaps [illegible]

Qy      [illegible] WKE[illegible]KMV[illegible]VVV[illegible]MEAM[illegible]WNPNVK[illegible]

Db      [illegible]WN[illegible]AV[illegible]YKVI[illegible]VVV[illegible]VMAAV[illegible]VWR[illegible]

Qy      [illegible]AAAA[illegible]VLA[illegible]MAI[illegible]

Db      [illegible]VMAA[illegible]

Qy      [illegible]

Db      [illegible]


RESULT [illegible]
[illegible]
hypothetical protein [illegible]
C;Species: [illegible]
C;Date: [illegible]
C;Accession: [illegible]
R;[illegible]
```

```
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, [illegible]; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic pathe
A;Reference number: A82950; MUID:20437[illegible]7
A;Accession: G83405
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1281 <STO>
A;Cross-references: GB:AE004618; GB:AE004[illegible]; NID:g9947[illegible]; PIDN:AAG[illegible]; GSPDB:GN00
A;Experimental source: strain PAO1
C;Genetics:
A;Gene: PA1923

  Query Match             6.8%;  Score [illegible];  DB 2;  Length 1281;
  Best Local Similarity   22.7%;  Pred. No. [illegible];
  Matches   61;  Conservative   39;  Mismatches   90;  Indels   79;  Gaps   [illegible];

Qy       32 CQEINWRALCDSCPC WMCQVRRR          [illegible]QEPALIYC[illegible]VAM [illegible]
             |  |  | |  :  |  ::  :|:|           |:               : :  :
Db     1042 CLSLAVPRLCCACPALYIALLRQE[illegible]AAL[illegible]LLYAS [illegible]

Qy       83 QKALGIINNQEGWKKESQQENGDEVLSKMVPDV[illegible]PVVVL[illegible]
             : | : || | :                          |: |  ||    |::  : |
Db     1102 LEMLRLANNLWGWQAAD                        [illegible]

Qy      142  AMGEW    NFNVKEIKVLQRIGKD                   [illegible]AAAAGNL  V 17[illegible]
              : ||    ||  : ::: |: :                      : |
Db     1141 LGLAEWPEEHNE[illegible]AKMALAI[illegible]

Qy      179 GPRDFVSVRCTKRRGSTCVLAGMATHPGEMR[illegible]VIRADHGPIC  MVLHILAGRPGK 23[illegible]
              |:|:                   ||  |||    :       :  |        |
Db     1200 TTREEFIE---- ---- -----RMACCPCEAAA[illegible]AAA[illegible]VAR[illegible]

Qy      236 IKLIW------ LLSIDLKGWLPKIIINQ 2[illegible]
                 |         ||::   | : ::::|:
Db     1246 EPQPWLRLAAAVLLALFAAGAVRQSVVNR 12[illegible]


RESULT  13
T25416
hypothetical protein T28D6.7 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-[illegible]-1999 #text_change 1[illegible]
C;Accession: T25416
R;Wilkinson, J.
submitted to the EMBL Data Library, October 1996
A;Reference number: Z20031
A;Accession: T25416
[illegible]atus: preliminary; translated from GB/EMBL/DDBJ
[illegible]lecule type: DNA
[illegible]esidues: 1-322 <WIL>
A;Cross-references: EMBL:Z81134; PIDN:CAB[illegible]; GSPDB:GN[illegible]; [illegible]T28D6.7
A;Experimental source: clone T28D6
C;Genetics:
A;Gene: CESP:T28D6.7
A;Map position: 3
A;Introns: 50/[illegible]; 127/[illegible]; 15[illegible]/[illegible]; 17[illegible]/3

  Query Match             6.7%;  Score 96;  DB 2;  Length 322;
  Best Local Similarity   23.0%;  Pred. No. [illegible];
  Matches   45;  Conservative   32;  Mismatches   76;  Indels   [illegible];  Gaps   [illegible];

Qy       83 QKALGIINNQEGW-----KKESQ--QENGDEVLS-MV      --PDVIKVEPIPVVVDQPM 12[illegible]
             :|    :   :  |:|       ||:       :| ::     :|:          ||       :::
Db       69 EFVKNICE[illegible]SECWVEVYKKKDITICTQN[illegible]MIKAIAQ[illegible]

Qy      130 DRLYEELVDRMEAMGEWNFNVKEIKVL[illegible]IVIIHHLAAAAAGNLV[illegible]
              ::: : : : :|   |||                :    : | : :     :  |   :
Db      129 EAEWDEYMIKQ[illegible]INEN              [illegible]
```

```
Qy      190 KRRGS[illegible]VLAGMA[illegible]MPEQS[illegible]VIRA[illegible]PICM[illegible]LAGSI SKIK  24[illegible]
                    ::          : |    : ||                       :: :: |  : |
Db      174 IEIDKDRI[illegible]SV[illegible]HEDYPPAKG[illegible]IRAIV [illegible]KIP 23[illegible]

Qy      239 IWLLSIDLKGWLPKII 254
             |||::    |     | |
Db      233 IWLVNRVIKVVALKVI 248


RESULT  14
D96611
probable CRK1 protein [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 02-Mar-2001
C;Accession: D96611
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; [illegible]; Kim, C.; White, O.; Alon
 Chin, C.W.; Chang, H.K.; Conn, L.; [illegible]; Tanaka, A.K.; Toriumi, [illegible]; Yewat,
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, [illegible]; Jenkins, [illegible]; Kim,
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; [illegible]; Matzia
Pizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, [illegible]; Schwartz, [illegible]; A.M.; Sun, H.; Tallo
Ker, M.; Wu, D.; Yu, [illegible]; Fraser, C.M.; [illegible]
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis
A;Reference number: A86141; MUID:2101[illegible]
A;Accession: D96611
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-686 <STO>
A;Cross-references: GB:AE005173; NID:g[illegible]; PIDN:AA[illegible]; GSPDB:GN00141
C;Genetics:
A;Gene: F8L23.17
A;Map position: 1


  Query Match             6.6%;  Score 91;  DB 2;  Length 686;
  Best Local Similarity   20.2%;  Pred. No. [illegible];
  Matches   47;  Conservative   32;  Mismatches   50;  Indels  [illegible]4;  Gaps   11;

Qy       74 YIQQGEVAMQ[illegible]NNQ[illegible]KKESQ[illegible]N[illegible]VLSKMVP[illegible]KVEPIPVVVDQPM 12[illegible]
             |: :    | :: |          | :  ||:|   |       |             :
Db       [illegible] YIETNHV[illegible]K       ENPKAGKKQ[illegible]VK-                          5[illegible]

Qy      1[illegible] DRLYEE[illegible]AMGEWNFNVKEIK VLQ [illegible]LAAA[illegible]         VGHEP 1[illegible]2
              |    : | :      | :  :|  : :|:                          |:      ||
Db       58 DTFSDEEVS[illegible]KKEESPEMKSCERVLQ[illegible]                     [illegible]NVGPL- 1[illegible]

Qy      183 FVSVRCTKR[illegible]STCVLAGMATHPGEM[illegible]CHV [illegible] LACGPGKIKLIWLL 242
                                           :   |          :    :|:  :|         |:
Db      1[illegible]                                QEEM[illegible]KAAK[illegible]MRG      WLV 1[illegible]

Qy      243 SI   DLKGWL[illegible]KII -   -    -    -    -   INQV[illegible]IEEAN 2[illegible]
             |:     :  | : ::                              ||:     :: | :
Db      129 SVACDAINGW[illegible]FSADSPEKLEEMIGQ[illegible]VTEA[illegible]LLINQ[illegible]KKVEEAN 1[illegible]


RESULT  15
T30225
polyketide synthase - Streptomyces hygroscopicus
C;Species: Streptomyces hygroscopicus
C;Date: 22-Oct-1999 #sequence_revision 2[illegible]-Oct-1999 #text_change 17-Nov-2000
C;Accession: T30225
R;Aparicio, J.F.; Molnar, I.; Schwecke, T.; [illegible]; Khaw, L.E.; St
Gene 169, 9-16, 1996
A;Title: Organization of the biosynthetic gene cluster for rapamycin in Streptomyces
A;Reference number: Z20782; MUID:9618[illegible]
A;Accession: T30225
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-10223 <APA>
A;Cross-references: EMBL:X86780; NID:g[illegible]; PIDN:CAA[illegible]
C;Genetics:
```

```
A;Gene: rapB
C;Superfamily: 3-oxoacyl-[acyl-carrier-protein] synthase [illegible]
C;Keywords: carrier protein
F;54-449/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase [illegible]
F;1533-1604/Domain: acyl carrier protein homology [illegible]
F;1637-2042/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase [illegible]
F;2135-2409/Domain: [acyl-carrier-protein] S-malonyltransferase [illegible]
F;3172-3243/Domain: acyl carrier protein homology [illegible]
F;3287-3681/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase [illegible]
F;3778-4052/Domain: [acyl-carrier-protein] S-malonyltransferase [illegible]
F;5160-5231/Domain: acyl carrier protein homology [illegible]
F;5275-5667/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase [illegible]
F;6776-6847/Domain: acyl carrier protein homology [illegible]
F;6891-7285/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase [illegible]
F;8411-8482/Domain: acyl carrier protein homology [illegible]
F;8526-8921/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase [illegible]
F;9012-9285/Domain: [acyl-carrier-protein] S-malonyltransferase [illegible]
F;[illegible]069-1011[illegible]/Domain: acyl carrier protein [illegible]

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches [illegible] Conservative [illegible] Mismatches [illegible]

Qy         24 PHQAVLAT[illegible]NWRA[illegible]WM[illegible]
Db       2085 K[illegible]AYLAA[illegible]

Qy         83 QKA[illegible]NN[illegible]WK[illegible]
Db       2141 [illegible]

Qy        134 EPLVDRMEAM[illegible]EWN[illegible]NVKEIKVI[illegible]
Db       2186 AVVDRVEVV[illegible]AAVW[illegible]

Qy        187 R[illegible]IKRRG[illegible]
Db       2245 VI[illegible]RSEATARG[illegible]MASVALPA[illegible]


Search completed: Aug [illegible]
Job time: 76 sec
```

```
                        GenCore version 4.5
              Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on:          August 22, 2001, 14:30:58 ; Search time 7.8 Seconds
                                            (without alignments)
                                            272.704 Million cell updates/sec

Title:           US-08-659-254-18
Perfect score:   1470
Sequence:        1 MLLATFKLCAGSSYRHMRNM ........... DFANHLRKRLESHPASEAR 285

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters:      93435

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       SwissProt_39:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                              SUMMARIES

                      %
Result              Query
   No.    Score     Match Length DB  ID                     Description
---------------------------------------------------------------------------
    1      1464     99.6    285  1  STAR_HUMAN              P49675 homo sapien
    2      1338     91.0    285  1  STAR_PIG                Q28996 sus scrofa
    3      1331     90.5    285  1  STAR_HORSE              O46689 equus cabal
    4    1305.5     88.8    284  1  STAR_RAT                P97826 rattus norv
    5      1299     88.4    285  1  STAR_BOVIN              Q28918 bos taurus
    6    1291.5     87.9    284  1  STAR_MOUSE              P51557 mus musculu
    7    1273.5     86.6    284  1  STAR_MESAU              P70114 mesocricetu
    8       536     36.5    121  1  STAR_SHEEP              P79245 ovis aries
    9       372     25.3    445  1  ML64_HUMAN              Q14849 homo sapien
   10       364     24.8    446  1  ML64_MOUSE              Q61542 mus musculu
   11     182.5     12.4    241  1  YX70_CAEEL              Q17883 caenorhabdi
   12       104      7.1    988  1  TNP6_ENTFA              [illegible] enterobacte
   13      94.5      6.4   3433  1  UTRO_HUMAN              P46939 homo sapien
   14        89      6.1    632  1  YKU4_YEAST              P36041 saccharomyc
   15        89      6.1   1018  1  YKC2_CAEEL              P41993 caenorhabdi
   16      88.5      6.0    422  1  TERL_BPSPP              P54[illegible] bacterioph
   17      88.5      6.0   1055  1  TIRB_HAEIN              [illegible] haemophilus
   18      87.5      6.0   1343  1  RPOB_HAEIN              P43738 haemophilus
   19        87      5.9    425  1  TIG_THEMA               Q9W[illegible] thermotoga
   20        87      5.9   1213  1  SMCX_MOUSE              P41230 mus musculu
   21      86.5      5.9    423  1  IDH_BACSU               P39126 bacillus su
   22      86.5      5.9   1225  1  Y309_MYCGE              P47551 mycoplasma
   23        86      5.9    218  1  YC35_METJA              Q27[illegible] methanococ
   24      85.5      5.8    295  1  FBX2_HUMAN              Q9UK22 homo sapien
   25      85.5      5.8    546  1  TCPQ_SCHPO              P78921 schizosacch
   26        85      5.8   1392  1  RPOB_NEIMA              P5[illegible] neisseria m
   27        85      5.8   1392  1  RPOB_NEIMB              Q9[illegible] neisseria m
   28        84      5.7    950  1  DC12_DROME              P18170 drosophila
   29        84      5.7   1123  1  DC11_DROME              P18[illegible] drosophila
   30        84      5.7   1589  1  DC13_DROME              P18171 drosophila
   31      83.5      5.7    420  1  Z229_HUMAN              Q9UJW7 homo sapien
   32      83.5      5.7    508  1  PDI2_ARATH              Q9SRG3 arabidopsis
   33        83      5.6    293  1  AR34_DROME              Q9VIM[illegible] drosophila
   34        83      5.6    401  1  CATE_ACICA              Q43935 acinetobact
   35        83      5.6    560  1  SYE_HAEWF               Q43794 haemophilus
   36        83      5.6    729  1  KAR4_YEAST              P17119 saccharomyc
   37      82.5      5.6    430  1  ENO_BACSU               P37869 bacillus su
   38      82.5      5.6    509  1  C931_SOYBN              Q42798 glycine max
   39      81.5      5.5    226  1  VATE_MANSE              P31402 manduca sex
   40      81.5      5.5    324  1  IDSA_METTM              Q58479 methanobact
   41        81      5.5    255  1  STX6_RAT                [illegible] rattus norv
   42        81      5.5    382  1  GMD2_CAEEL              Q45183 caenorhabdi
   43        81      5.5   1263  1  RPOB_THEMA              P29398 thermotoga
   44        81      5.5   1560  1  SMCX_HUMAN              P41229 homo sapien
   45        81      5.5   2032  1  CTO6_HUMAN              Q14648 homo sapien

                              ALIGNMENTS

RESULT    1
STAR_HUMAN
ID   STAR_HUMAN     STANDARD;      PRT;   285 AA.
AC   P49675; Q16396;
DT   01-FEB-1996 (Rel. 33, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN PRECURSOR.
GN   STAR.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID 9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Adrenal cortex;
RX   MEDLINE 95281549; PubMed 7761400;
RA   Sugawara T., Holt J.A., Driscoll D., Strauss J.F. III, Lin D.,
RA   Miller W.L., Patterson D., Clancy K.P., Hart I.M., Clark B.J.,
RA   Stocco D.M.;
RT   "Human steroidogenic acute regulatory protein: functional activity in
RT   COS-1 cells, tissue-specific expression, and mapping of the
RT   structural gene to 8p11.2 and a pseudogene to chromosome 13.";
RL   Proc. Natl. Acad. Sci. U.S.A. 92:4778-4782(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Placenta;
RX   MEDLINE 96038218; PubMed 7547998;
RA   Sugawara T., Lin D., Holt J.A., Martin K.O., Javitt N.B., Miller W.L.,
RA   Strauss J.F. III;
RT   "Structure of the human steroidogenic acute regulatory protein (StAR)
RT   gene: StAR stimulates mitochondrial cholesterol 27-hydroxylase
RT   activity.";
RL   Biochemistry 34:12506-12512(1995).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96008588; PubMed 7548190;
RA   Gradi A., Tang Wai R., McBride H.M., Chu L.L., Shore G.C.,
RA   Pelletier J.;
RT   "The human steroidogenic acute regulatory (StAR) gene is expressed in
RT   the urogenital system and encodes a mitochondrial polypeptide.";
RL   Biochim. Biophys. Acta 1258:228-233(1995).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RA   Yu W., Sarginson J., Gibbs R.A.;
RL   Submitted (NOV-1997) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   VARIANT CLAH LEU-182.
RX   MEDLINE 97056837; PubMed 8948562;
RA   Bose H.S., Sugawara T., Strauss J.F. III, Miller W.L.;
RT   "The pathophysiology and genetics of congenital lipoid adrenal
RT   hyperplasia.";
RL   New Engl. J. Med. 335:1870-1878(1996).
RN   [6]
RP   VARIANT ALA 2--
```

```
RX   MEDLINE 98[illegible] PubMed [illegible]
RA   Katsumata N., [illegible] Nakazawa [illegible]
RA   Shimizu M., [illegible]
RT   "A novel frameshift mutation [illegible] and [illegible]
RT   the steroidogenic acute regulatory protein [illegible]
RT   with congenital lipoid adrenal hyperplasia."
RL   Hum. Mutat. [illegible]
CC   -!- FUNCTION: PLAYS A KEY ROLE IN STEROID HORMONE [illegible]
CC       ENHANCING THE METABOLISM OF [illegible]
CC       TRANSPORTER THAT BINDS [illegible] AND TRANS[illegible]
CC       INTERMEMBRANE SPACE OF THE MITOCHONDRION [illegible]
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN GONADS [illegible]
CC       KIDNEY.
CC   -!- DISEASE: DEFECTS IN STAR CAUSE CONGENITAL [illegible]
CC       HYPERPLASIA (LA[illegible] [illegible]
CC   -!- SIMILARITY: BELONGS TO THE STAR FAMILY.
CC
CC   This SWISS-PROT entry is copyright. [illegible]
CC   between the Swiss Institute of Bioinformatics [illegible] EMBL [illegible]
CC   the European Bioinformatics Institute. [illegible]
CC   use by non-profit institutions [illegible]
CC   modified and this statement is not removed. [illegible]
CC   entities requires a license agreement [illegible]
CC   or send an email to license@isb-sib.ch).
CC
DR   EMBL; [illegible]28[illegible]; AA[illegible]
DR   EMBL; [illegible]10[illegible]; AA[illegible]
DR   EMBL; [illegible]099[illegible]; AA[illegible]; JOINED.
DR   EMBL; U29100; AA[illegible]; JOINED.
DR   EMBL; U29101; AA[illegible]; JOINED.
DR   EMBL; U29102; AA[illegible]; JOINED.
DR   EMBL; U29103; AA[illegible]; JOINED.
DR   EMBL; U29104; AA[illegible]; JOINED.
DR   EMBL; [illegible]669; AA[illegible]
DR   EMBL; AF035[illegible]; AAB8[illegible]
DR   MIM; 6[illegible]617; -.
DR   MIM; 201710; -.
DR   InterPro; IPR[illegible];
DR   InterPro; IPR[illegible];
DR   Pfam; PF[illegible]; START; [illegible]
DR   PRINTS; PR00978; STARPROTEIN.
KW   Lipid binding; Lipid transport; [illegible]
KW   Mitochondrion; Transit peptide; [illegible]
FT   TRANSIT      [illegible]                 MITOCHONDRION [illegible]
FT   CHAIN        [illegible]    285          STEROIDOGENIC ACUTE [illegible]
FT   VARIANT      [illegible]    18[illegible]        [illegible] (IN [illegible])
FT                                          /FTId=VAR_[illegible]
FT   VARIANT      2[illegible]   2[illegible]          [illegible] A
FT                                          /FTId=VAR_[illegible]
SQ   SEQUENCE   285 AA;  [illegible] MW;  [illegible]

  Query Match             [illegible]%;  Score [illegible];  DB [illegible];
  Best Local Similarity   [illegible]%;  Pred. No. [illegible];
  Matches  2[illegible];  Conservative [illegible];  Mismatches [illegible];  [illegible]

Qy          1 MLLATFKL[illegible]YRHMRNMK[illegible]LAVMA[illegible]N[illegible]
              [illegible]
Db          1 MLLATFKL[illegible]YRHMRNMK[illegible]LAVMA[illegible]N[illegible]

Qy         61 SFLEFTLY[illegible]AY[illegible]FAMQKA[illegible]N[illegible]WRR[illegible]
              [illegible]
Db         61 SFLEFTLY[illegible]AY[illegible]FAMQKA[illegible]N[illegible]WRR[illegible]

Qy        121 RLEVVVDQPMER[illegible]YEE[illegible]MEAM[illegible]WN[illegible]NVK[illegible]
              [illegible]
Db        121 RLEVVVDQPMER[illegible]YEE[illegible]MEAM[illegible]WN[illegible]NVK[illegible]

Qy        181 PRDFVSVRCAKR[illegible]SS[illegible]A[illegible]MA[illegible]NM[illegible]
              [illegible]
Db        181 PRDFVSVRCAKR[illegible]SS[illegible]A[illegible]M[illegible]NM[illegible]
```

```
Qy        [illegible] W[illegible] K W[illegible] N[illegible] V[illegible] LVLEAN[illegible] A[illegible]
              [illegible]
Db        [illegible] W[illegible] K W[illegible] N[illegible] V[illegible] LVLEAN[illegible] A[illegible]


RESULT [illegible]
STAR_[illegible]
ID   STAR_[illegible]     STANDARD;      PRT;   2[illegible] AA.
AC   [illegible]
DT   [illegible] NOV [illegible] (Rel. [illegible], Created)
DT   [illegible] NOV [illegible] (Rel. [illegible], Last sequence update)
DT   [illegible] MAY [illegible] (Rel. [illegible], Last annotation update)
DE   [illegible] ACUTE REGULATORY PROTEIN [illegible]
GN   STAR.
OS   [illegible]
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; [illegible]
OC   Mammalia; [illegible]
OX   NCBI_TaxID=[illegible]
RN   [1]
RP   SEQUENCE FROM N.A.
RC   [illegible]
RX   MEDLINE [illegible] PubMed [illegible]
RA   [illegible] N[illegible]
RA   [illegible]
RT   "[illegible] (STAR
RT   [illegible]".
RL   [illegible]
RN   [2]
RP   SEQUENCE [illegible] FROM N.A.
RC   [illegible]
RX   MEDLINE [illegible] PubMed [illegible]
RA   [illegible]
RA   [illegible]
RT   "[illegible]
RT   protein (StAR) [illegible] with
RT   [illegible]".
RL   [illegible]
CC   -!- FUNCTION: PLAYS A KEY ROLE IN STEROID HORMONE SYNTHESIS BY
CC       ENHANCING THE METABOLISM OF CHOLESTEROL INTO PREGNENOLONE.
CC       TRANSPORTER THAT BINDS [illegible] AND TRANSFERS CHOLESTEROL [illegible]
CC       INTERMEMBRANE SPACE OF THE MITOCHONDRION [illegible]
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL.
CC   -!- [illegible] IN [illegible]
CC       [illegible] ADRENAL
CC       [illegible] AS WELL AS [illegible] PREGNANCY [illegible]
CC       DEVELOPING [illegible] AND PREGNANCY [illegible] IN
CC       THE [illegible]
CC   -!- SIMILARITY: BELONGS TO THE STAR FAMILY.
CC
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC
DR   EMBL; [illegible]; AAB[illegible]
DR   EMBL; [illegible]; AAB[illegible]
DR   [illegible]
DR   InterPro; [illegible]
DR   [illegible]
DR   PRINTS; PR[illegible]; [illegible]
KW   [illegible]
KW   Mitochondrion; [illegible] peptide.
FT   TRANSIT      [illegible]                 MITOCHONDRION [illegible]
FT   CHAIN        [illegible]                 STEROIDOGENIC ACUTE REGULATORY PROTEIN.
FT   [illegible]
SQ   SEQUENCE   [illegible] AA;  [illegible] MW;  [illegible]

  Query Match             [illegible]
  Best Local Similarity   [illegible]  Pred. No. [illegible]
  Matches  254;  Conservative   18;  Mismatches   13;  Indels    [illegible];  Gaps   [illegible]

Qy        1 MLLATFKLCACSSYRHMRNMKGLRQQAVMAISQELNRRALGGPTPS[illegible]WINQVRRRSSLL[illegible] [illegible]
            |||||||||||||||:||||||| |||: : |||:|||||||||   :||||||||||||
Db        1 MLLATFKLCA[illegible]YRHVRNMK[illegible]LSHQAVI[illegible] [illegible]

Qy       61 CFLEFTLYS[illegible]AY[illegible]FAMQRAL[illegible] [illegible]
            |:||:| ||||:|||:|||||||||:|| | ||||||||||:|:||||:|:||||:|||||
Db       61 SQLEDTFYSDQDLAYIQQ[illegible]FAMQFALDI[illegible]SNL[illegible] [illegible]

Qy      121 R[illegible]EVVVDQPMPR[illegible]YPF[illegible]VPRMPAMGPWNPNVKF[illegible] [illegible]
            ||||||||||||||||||||||||||||||||:|||: |:||||||| |||:||||:| |  |
Db      121 R[illegible]EVVVD[illegible]PMP[illegible]YPF[illegible]VF[illegible]MEAMGPWNP[illegible]VK[illegible] [illegible]

Qy      181 PRDFV[illegible]VRCAKRRGSTCVLAGMATDFGNM[illegible]RAP[illegible]PT[illegible]MVLHPLA[illegible] [illegible]
            ||||||| | ||||| ||||||||||| |||||| ||||||||||||:||||||||||||
Db      181 PRDFV[illegible]VC[illegible]KR[illegible]GSVCVLAGMATDFGNM[illegible] [illegible]

Qy      241 WLLSIPLKCWLPKS[illegible]INQVLSQTQVLFANGLPR[illegible]AGLAR[illegible] [illegible]
            |||||||||||||:||||||||||||||||||||||  |  :  :|||
Db      241 WLLSIDLKCWLPKT[illegible]INQVLSQTQVDFAN[illegible] [illegible]

[illegible]LT    3
[illegible]_HORSE
ID   STAR_HORSE      STANDARD;      PRT;   [illegible] AA.
AC   O46689;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN PRECURSOR.
GN   STAR.
OS   Equus caballus (Horse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX   NCBI_TaxID=9796;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Follicle;
RX   MEDLINE=99124482; PubMed=9927292;
RA   Kerban A., Boerboom D., Sirois J.;
RT   "Human chorionic gonadotropin induces an inverse regulation of
RT   steroidogenic acute regulatory protein messenger ribonucleic acid in
RT   theca interna and granulosa cells of equine preovulatory follicles.";
RL   Endocrinology 140:667-674(1999).
CC   -!- FUNCTION: PLAYS A KEY ROLE IN STEROID HORMONE SYNTHESIS BY
CC       ENHANCING THE METABOLISM OF CHOLESTEROL INTO PREGNENOLONE
CC       TRANSPORTER THAT BINDS TO AND TRANSPORT CHOLESTEROL [illegible] THE
CC       INTERMEMBRANE SPACE OF THE MITOCHONDRION (PROBABLE).
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL
CC   -!- SIMILARITY: BELONGS TO THE STAR FAMILY.
[illegible]   --------------------------------------------------------------------------
[illegible]   This SWISS-PROT entry is copyright. It is produced through a collaboration
[illegible]   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://[illegible]
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AF031697; AAC04704.1; -.
DR   EMBL; AF031696; AAC04703.1; -.
DR   InterPro; IPR000799; -.
DR   InterPro; IPR002913; -.
DR   Pfam; PF01852; START; 1.
DR   PRINTS; PR00978; STARPROTEIN.
KW   Lipid-binding; Lipid transport; Transport; Steroidogenesis;
KW   Mitochondrion; Transit peptide.
FT   TRANSIT       1      [illegible]       MITOCHONDRION (POTENTIAL).
FT   CHAIN         [illegible]    285       STEROIDOGENIC ACUTE REGULATORY PROTEIN.
SQ   SEQUENCE   285 AA;  31852 MW;  B[illegible]9727D CRC64;
```

```
  Query Match             90.5%;  Score [illegible];  DB [illegible];  Length 285;
  Best Local Similarity   90.5%;  Pred. No. 2.3e-97;
  Matches  258;  Conservative   11;  Mismatches   16;  Indels    [illegible];  Gaps    0;

Qy        1 MLLA[illegible]HMRNMKGLRQQAVM[illegible] [illegible] 60
            |||||||||| ||||:||||||| |  :|   ||| |:  |||     ||||:||||
Db        1 MLLA[illegible]RLVRNMKGLRQ[illegible] [illegible] 60

Qy       61 SRLEEILY[illegible]AY[illegible]EAMQKAL[illegible]VKRVVEDVGRVF 120
            |:||:|||||||||||:|||||||||||||||   |  |||  |:|:| |     |:|||||||
Db       61 [illegible]QIEE[illegible]AMQKA[illegible]VGKVF 120

Qy      121 R[illegible]VVV[illegible]VLRMEAMGEWN[illegible] [illegible] 180
            |||| |||||:|:|||||||||||||||         |      |||||: | ||||:||:||||| |
Db      121 R[illegible]VEVL[illegible]VERMEAMGEWN[illegible] [illegible]NLVG 180

Qy      181 PRDFVSV[illegible]VLAGMATDF[illegible] [illegible]GKIKL 240
            ||||||||||  ||||||||||||| |  |     |    ||:||      |:||||:||||  |
Db      181 PRDFVNVR[illegible]VLA[illegible] [illegible]KIKT 240

Qy      241 WLLS[illegible]INQVLS[illegible]APEAP[illegible] 285
            ||||||||:|:||||||||||||||||  :|  |   |    |    |:
Db      241 WLLSIDLKCWLPKTIINQVLSQTQVDFANH[illegible]APEAP[illegible] 285

RESULT    4
STAR_RAT
ID   STAR_RAT       STANDARD;      PRT;   285 AA.
AC   P97826; P97628;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN PRECURSOR.
GN   STAR.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=WISTAR; TISSUE=Ovary;
RX   MEDLINE=97466899; PubMed=9328229;
RA   Mizutani T., Sonoda Y., Minegishi T., Wakabayashi K., Miyamoto K.;
RT   "Molecular cloning, characterization and cellular distribution of rat
RT   steroidogenic acute regulatory protein (StAR) in the ovary.";
RL   Life Sci. 61:1497-1506(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY; TISSUE=Testis;
RX   MEDLINE=97167237; PubMed=9015375;
RA   Lee H.K., Ahn R.S., Kwon H.B., Soh J.;
RT   "Nucleotide sequence of rat steroidogenic acute regulatory protein
RT   complementary DNA.";
RL   Biochem. Biophys. Res. Commun. 23[illegible](1997)
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY; TISSUE=Adrenal gland;
RX   MEDLINE=98184371; PubMed=9516465;
RA   Ariyoshi N., Kim Y.-C., Artemenko I., Bhattacharyya K.K.,
RA   Jefcoate C.R.;
RT   "Characterization of the rat Star gene that encodes the predominant
RT   3.5-kilobase pair mRNA. ACTH stimulation of adrenal steroids in vivo
RT   precedes elevation of Star mRNA and protein.";
RL   J. Biol. Chem. 273:7610-7619(1998).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY; TISSUE=Adrenal gland;
RX   MEDLINE=9904896[illegible]; PubMed=983213[illegible];
RA   Furukawa A., Miyatake A., Ohnishi T., Ichikawa Y.;
RT   "Steroidogenic acute regulatory protein (StAR) transcripts
RT   constitutively expressed in the adult rat central nervous system:
RT   colocalization of StAR, cytochrome P-450scc (CYP XIA1), and 3beta
RT   hydroxysteroid dehydrogenase in the rat brain.";
```

```
RL   J. Neurochem. [illegible]
CC   -!- FUNCTION: PLAYS A KEY ROLE IN [illegible] NE [illegible]
CC       ENHANCING THE METABOLISM [illegible]
CC       TRANSPORTER THAT BINDS TO AND TRANSPORT [illegible]
CC       INTERMEMBRANE SPACE [illegible] THE M[illegible]
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL
CC   -!- SIMILARITY: [illegible] THE STAR FAMILY.
CC   --------
CC   This SWISS-PROT entry [illegible]
CC   between the Swiss Institute [illegible]
CC   the European Bioinformatics [illegible]
CC   use by non-profit institutions [illegible]
CC   modified and this statement [illegible]
CC   entities requires a license agreement [illegible]
CC   or send an email to [illegible]
CC   --------
DR   EMBL; AB001[illegible]; BAA[illegible]
DR   EMBL; [illegible]641[illegible]; AA[illegible]
DR   EMBL; AF044[illegible]; AAC[illegible]
DR   EMBL; AB006[illegible]; BAA[illegible]
DR   InterPro; IPR[illegible]
DR   InterPro; IPR[illegible]
DR   Pfam; PF0185[illegible]; START; [illegible]
DR   PRINTS; PR00[illegible]; STARPROTEIN.
KW   Lipid-binding; Lipid transport; [illegible]
KW   Mitochondrion; Transit peptide.
FT   TRANSIT      1    [illegible]       MITOCHONDRION [illegible]
FT   CHAIN        [illegible]  28[illegible]       STEROIDOGENIC ACUTE [illegible]
FT   CONFLICT     [illegible]                MLLATFK[illegible]
FT                                           [illegible]
FT                                           [illegible]
SQ   SEQUENCE   284 AA;  [illegible] MW;  [illegible]

  Query Match             [illegible]
  Best Local Similarity   [illegible]
  Matches  [illegible];  Conservative  [illegible];  Mismatches  [illegible]

Qy          1 MLLATFKL[illegible]AG[illegible]YRHMRNMK[illegible]LP[illegible]AVMA[illegible]NERAL[illegible]
Db          1 MLLATFKL[illegible]AG[illegible]YRHMRNMK[illegible]LPH[illegible]AVLA[illegible]NERAL[illegible]

Qy         61 SPLEETLY[illegible]AYL[illegible]GEEAMQKAL[illegible]NQ[illegible]WEKL[illegible]
Db         60 SQLEATLYS[illegible]SYL[illegible]GEEAMQKAL[illegible]NNQ[illegible]WEKL[illegible]

Qy        121 R[illegible]EVVVD[illegible]EM[illegible]YEL[illegible]MEAM[illegible]WN[illegible]NVK[illegible]
Db        120 R[illegible]EVLLD[illegible]EM[illegible]YEL[illegible]MEAM[illegible]WN[illegible]NVK[illegible]

Qy        181 PR[illegible]VSVK[illegible]AR[illegible]VIA[illegible]MA[illegible]NM[illegible]
Db        180 PR[illegible]VSVK[illegible]VIA[illegible]MA[illegible]

Qy        241 WL[illegible]IDLK[illegible]WL[illegible]VL[illegible]A[illegible]
Db        240 WL[illegible]IDLK[illegible]WL[illegible]VL[illegible]A[illegible]

RESULT 5
STAR_BOVIN
ID   STAR_BOVIN     STANDARD;      PRT;   [illegible] AA.
AC   Q28918;
DT   01-NOV-1997 (Rel. [illegible], Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   30-MAY-2000 (Rel. [illegible], Last annotation update)
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN [illegible]
GN   STAR.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; [illegible]
OC   Mammalia; Eutheria; [illegible]
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
```

```
RN   [illegible]
RP   [illegible]
RX   MEDLINE[illegible]
RA   [illegible]
RT   [illegible]
RT   [illegible]
RL   [illegible]
RN   [illegible]
RP   [illegible]
RA   [illegible]
RL   Submitted (MAY-[illegible]) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: PLAYS A KEY ROLE IN [illegible] SYNTHESIS BY
CC       ENHANCING THE METABOLISM [illegible] PREGNENOLONE.
CC       TRANSPORTER THAT BINDS [illegible] AND TRANSPORT [illegible]
CC       INTERMEMBRANE SPACE OF THE M[illegible] PARTS.
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL
CC   -!- [illegible] HUMAN [illegible] ADRENAL GLAND
CC   -!- SIMILARITY: [illegible] THE STAR FAMILY.
CC   --------
CC   This SWISS-PROT entry is copyright. [illegible] collaboration
CC   between [illegible] EMBL [illegible]
CC   the European Bioinformatics Institute. [illegible]
CC   [illegible] was
CC   modified [illegible] commercial
CC   entities [illegible]
CC   [illegible]
CC   --------
DR   EMBL; [illegible]; AAA[illegible]
DR   EMBL; [illegible]; AAA[illegible]
DR   [illegible]
DR   InterPro; IPR[illegible]
DR   [illegible]
DR   PRINTS; PR00[illegible]; STARPROTEIN.
KW   [illegible]
KW   Mitochondrion; Transit peptide.
FT   TRANSIT      [illegible]       MITOCHONDRION (POTENTIAL).
FT   CHAIN        [illegible]       STEROIDOGENIC ACUTE REGULATORY PROTEIN.
FT   CONFLICT     [illegible]       [illegible] Y (IN [illegible] SEQUENCE).
FT   CONFLICT     [illegible]       [illegible] K (IN [illegible] SEQUENCE).
SQ   SEQUENCE   [illegible] AA;  [illegible] MW;  [illegible]

  Query Match             [illegible]
  Best Local Similarity   [illegible]
  Matches  [illegible];  Conservative  [illegible];  Mismatches  [illegible]

Qy          1 M[illegible]AVMA[illegible]WN[illegible]VRRR[illegible]
Db          1 M[illegible]AVLA[illegible]AAWINLVRRR[illegible]

Qy         [illegible] EAMQKA[illegible]WEK[illegible]KVMS[illegible]VVEV[illegible]KV[illegible]
Db         [illegible] EAMQKA[illegible]WEK[illegible]ANSLVLSKV[illegible]KV[illegible]

Qy        [illegible] VVVL[illegible]VEGMEAM[illegible]WNENVKE[illegible]EV[illegible]LHRLAAPAA[illegible]NLV[illegible]
Db        [illegible] AVVL[illegible]VE[illegible]AM[illegible]WNENVKE[illegible]EV[illegible]LHRLAAEVAGNLV[illegible]

Qy        [illegible] VSVK[illegible]AR[illegible]MA[illegible]NMI[illegible]VI[illegible]
Db        [illegible] MA[illegible]NMI[illegible]VI[illegible]

Qy        [illegible] ANE[illegible]
Db        [illegible] ANE[illegible]

RESULT [illegible]
STAR_M[illegible]
ID   STAR_M[illegible]   STANDARD;   [illegible] AA.
AC   [illegible]
DT   [illegible], Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN PRECURSOR (LUTEINIZING HORMONE
DE   INDUCED PROTEIN).
GN   STAR.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID 10090;
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX   MEDLINE=95050616; PubMed 7961770;
RA   Clark B.J., Wells J., King S.R., Stocco D.M.;
RT   "The purification, cloning, and expression of a novel luteinizing
RT   hormone-induced mitochondrial protein in MA-10 mouse Leydig tumor
RT   cells. Characterization of the steroidogenic acute regulatory protein
RT   (StAR).";
RL   J. Biol. Chem. 269:28314-28322(1994).
CC   -!- FUNCTION: PLAYS A KEY ROLE IN STEROID HORMONE SYNTHESIS BY
CC       ENHANCING THE METABOLISM OF CHOLESTEROL INTO PREGNENOLONE.
CC       TRANSPORTER THAT BINDS TO AND TRANSPORT CHOLESTEROL THR[illegible] THE
CC       INTERMEMBRANE SPACE OF THE MITOCHONDRION (PROBABLE).
     -!- SUBCELLULAR LOCATION: MITOCHONDRIAL.
     -!- SIMILARITY: BELONGS TO THE STAR FAMILY.
     -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as  long as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement [illegible]
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; L36062; AA[illegible].1; -.
DR   MGD; MGI:102760; Star.
DR   InterPro; IPR000799; -.
DR   InterPro; IPR002913; -.
DR   Pfam; PF01852; START; 1.
DR   PRINTS; PR00978; STARPROTEIN.
KW   Lipid-binding; Lipid transport; Transport; Steroidogenesis;
KW   Mitochondrion; Transit peptide.
FT   TRANSIT       1      ?       MITOCHONDRION (POTENTIAL).
FT   CHAIN         ?    284       STEROIDOGENIC ACUTE REGULATORY PROTEIN.
FT   DOMAIN      170    174       POLY-ALA.
SQ   SEQUENCE   284 AA;  [illegible]1626 MW;  [illegible] CRC64;


  Query Match             87.9%;  Score [illegible];  DB 1;  Length 284;
  Best Local Similarity   87.0%;  Pred. No. 2.8e-94;
  Matches  248;  Conservative   19;  Mismatches   17;  Indels   [illegible];  Gaps   [illegible];

            1 MLLATFKLCAGSSYRHMRNMKGLRQQAVMA[illegible]NEEALGGPTPCHINLVERRSSLL[illegible]
              | |||||||||:||||||||||||| ||:::  | ||   |||   :|  |:  ||  | ||||
Db          1 MFLATFKLCAGSSYRHMRNMKGLRHQAVLA[illegible]NWRALGDSSPG[illegible]WMGQVRRRSSLL[illegible]

Qy         61 SRLEETLYSDQELAYLQQGEEAMQKAL[illegible]GWKKESQQDNGEKVMSKVVPDVGKVF [illegible]
              |:|| ||||||||:|:||||  ||||||||||:|   ||.||||||:|||:|:|||:  |||:||||
Db         60 SQLEATLYSDQELSYIQQGEVAMQKAL[illegible]GWKKESQQENG[illegible]EM[illegible]KVF [illegible]

Qy        121 RLEVVVDQPMERLYEELVERMEAMGEWNPNVKEIKVLQKIGKDTFITHELAAEAAGNLVG [illegible]
              ||||||||||:||||||||:|||||||||||||  |.|||||:||||| ||||||| ||||||
Db        120 RLEVVVDQPMDRLYEELVDRMEAMGEWNPNVKEIKV[illegible]KDTVITHELAAAAAGNLV[illegible]

Qy        181 PRDFVSVRCAKRRGSTCVLAGMATDFGNMPEQK[illegible]VIRAEHGPTCMVLHPLAGSPSKTKLT [illegible]
              |||||:||| ||||||||||||||||| |'  |:|     |||||||||||||:|:|:| | :|||:|
Db        180 PRDFVSVRCIKRRGSTCVLAGMATHFGEMPEQ[illegible]GVIRAEHGPTCMVLHPLAG[illegible]SKTKLT [illegible]

Qy        241 WLLSIDLKGWLPKSIINQVLSQTQVDFA[illegible]HLRKRLESHPASEARC 28[illegible]
              |||||||||||||:||||||||||||::   ||  |  |  |:  ||||:|
Db        240 WLLSIDLKGWLPKSIINQVLSQTQIDFANHLRKRLEASPASEAQC 284

RESULT    7
```

```
STAR_MESAU
ID   STAR_MESAU     STANDARD;      PRT;   284 AA.
AC   P70114;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN PRECURSOR.
GN   STAR.
OS   Mesocricetus auratus (Golden hamster).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC   Mesocricetus.
OX   NCBI_TaxID [illegible];
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=9712[illegible]; PubMed [illegible];
RA   Fleury A., Cloutier M., Ducharme L., Lefebvre A., Lehoux J.,
RA   Lehoux J.-G.;
RT   "Adrenocorticotropin regulates the level of the steroidogenic acute
RT   regulatory (StAR) protein mRNA in hamster adrenals.";
RL   Endocr. Res. [illegible](1996).
CC   -!- FUNCTION: PLAYS A KEY ROLE IN STEROID HORMONE SYNTHESIS BY
CC       ENHANCING THE METABOLISM OF CHOLESTEROL INTO PREGNENOLONE.
CC       TRANSPORTER THAT BINDS TO AND TRANSPORT CHOLESTEROL THROUGH THE
CC       INTERMEMBRANE SPACE OF THE MITOCHONDRION (PROBABLE).
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL.
CC   -!- SIMILARITY: BELONGS TO THE STAR FAMILY.
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as  long as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; U66490; AAB06763.1; -.
DR   InterPro; IPR000799; -.
DR   InterPro; IPR002913; -.
DR   Pfam; PF01852; START; 1.
DR   PRINTS; PR00978; STARPROTEIN.
KW   Lipid-binding; Lipid transport; Transport; Steroidogenesis;
KW   Mitochondrion; Transit peptide.
FT   TRANSIT       1      ?       MITOCHONDRION (POTENTIAL).
FT   CHAIN         ?    284       STEROIDOGENIC ACUTE REGULATORY PROTEIN.
SQ   SEQUENCE   284 AA;  31497 MW;  [illegible] CRC64;


  Query Match             86.6%;  Score [illegible];  DB [illegible];  Length 284;
  Best Local Similarity   86.7%;  Pred. No. [illegible];
  Matches  247;  Conservative   17;  Mismatches   20;  Indels    1;  Gaps    1;

Qy          1 MLLATFKLCA[illegible]SYRHMRNMKGLRQ[illegible]NLVERRSSLL[illegible]
              |||||||||||||:|||||||| :|     |  |  :     ||:||||||| |
Db          1 MLLATFKLCAGSSYRHLRNMKGLR[illegible]WIGQVRRRSSLLG 59

Qy         61 SRLEETLYSDQELAYLQQGEEAMQKAL[illegible]NQ[illegible]GWKKESQQ[illegible]GEKVMSKVVPDVGKVF [illegible]
              |:|:  ||:|||:|:|||||||||||   |  |  : :  ||:|:|||||| ||
Db         60 SQLEAALYSDQELSYIQQGEVAMQKAL[illegible]NQ[illegible]GWKKENQ[illegible]N[illegible]EVLSKVVPDVGKVF 119

Qy        121 RLEVVVDQPMERLYEELVERMEAMGEWN[illegible]NVKE[illegible]KVLQKIGK[illegible]HELAAEAAGNLVG 180
              |:|||  | |:|| |||:|||||   | |  | |  :||   :|||||| |||||||
Db        120 RLEVVVDQP[illegible]LYAELVDRMEAMGEWN[illegible]NVKE[illegible]K[illegible]VITHELAAAAAGNLVG 179

Qy        181 PRDFVSVRCAKRRGSTCVLAGMATDFGNMPEQKGVIRAEHGPTCMVLHPLAGSPSKTKLT 240
              |||:|||| |||||||||||:|:|  |    |  |  ||  :|:| |||:||||||:|
Db        180 PRDFVSVRCAKRRGSTCVLAGMATHFGNMPEQKGVIRAEHGPTCMVLHPLAGSPSKTKLT 239

Qy        241 WLLSIDLKGWLPKSIINQVLSQTQVDFANHLRKRLESHPASEARC 285
              ||||  |:||:|||||||||||::|:   |||||    ||
Db        240 WLLSIDLKGWLPKSIINQVLSQTQMEFANHLRKRLESSSASEARC 284
```

```
RESULT [illegible]
STAR_SHEEP
ID   STAR_SHEEP     STANDARD;      PRT;   [illegible] AA.
AC   [illegible]
DT   30-MAY-2000 (Rel. [illegible], Created)
DT   30-MAY-2000 (Rel. [illegible], Last sequence update)
DT   30-MAY-2000 (Rel. [illegible], Last annotation update)
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN (FRAGMENT).
GN   STAR.
OS   Ovis aries (Sheep).
OC   Eukaryota; Metazoa; Chordata; [illegible]
OC   Mammalia; Eutheria; [illegible]
OC   Bovidae; Caprinae; [illegible]
OX   NCBI_TaxID=[illegible]
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Corpus luteum;
RX   MEDLINE=[illegible]; PubMed=[illegible];
RA   Juengel J.L., Meberg B.M., Turzillo A.M., Nett T.M., [illegible]
RT   "Hormonal regulation of messenger ribonucleic acid [illegible]
RT   steroidogenic acute regulatory protein [illegible]"
RL   Endocrinology [illegible](1995).
CC   -!- FUNCTION: PLAYS A KEY ROLE IN STEROID HORMONE [illegible] BY
CC       ENHANCING THE METABOLISM OF CHOLESTEROL [illegible]
CC       TRANSPORTER THAT BINDS [illegible] AND TRANSFERS [illegible]
CC       INTERMEMBRANE SPACE OF THE MITOCHONDRION [illegible]
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL.
CC   -!- SIMILARITY: BELONGS TO THE STAR FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. [illegible]
CC   between the Swiss Institute [illegible]
CC   the European Bioinformatics Institute. [illegible]
CC   use by non-profit institutions [illegible]
CC   modified and this statement [illegible]
CC   entities requires a license agreement [illegible]
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; [illegible]; AA[illegible]
DR   InterPro; IPR[illegible]
DR   PRINTS; PR[illegible]; STARPROTEIN.
KW   Lipid-binding; [illegible]
KW   Mitochondrion.
FT   NON_TER [illegible]
SQ   SEQUENCE   [illegible] AA;   [illegible] MW;   [illegible] CRC64;

  Query Match        [illegible]
  Best Local Similarity   [illegible]
  Matches [illegible]; Conservative [illegible]; Mismatches [illegible]

Qy        97 KKLSQQDN[illegible]KVMSKV[illegible]KVEE[illegible]VVV[illegible]
Db         1 KKENRQANS[illegible]KV[illegible]KVEE[illegible]VVV[illegible]

Qy       157 L[illegible]KLGKI[illegible]AA[illegible]AGNLV[illegible]
Db        61 L[illegible]GKI[illegible]AA[illegible]AGNLV[illegible]

Qy       217 K 217
             |
Db       121 K 121


RESULT [illegible]
ML64_HUMAN
ID   ML64_HUMAN     STANDARD;      PRT;   [illegible] AA.
AC   [illegible]
DT   [illegible] MAY 2000 (Rel. [illegible], Created)
DT   [illegible] MAY 2000 (Rel. [illegible], Last sequence update)
DT   [illegible] (Rel. [illegible], Last annotation update)
DE   MLN 64 PROTEIN [illegible]
GN   MLN64 OR CAB1.
OS   Homo sapiens (Human).
```

```
OC   Eukaryota; Metazoa; Chordata; [illegible] Vertebrata; [illegible]
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=[illegible]
RN   [1]
RP   SEQUENCE FROM N.A.
RC   [illegible]
RX   MEDLINE=[illegible]; PubMed=[illegible];
RA   [illegible]
RA   [illegible]
RT   "[illegible]
RT   [illegible]
RT   [illegible]"
RL   [illegible]
RN   [2]
RP   SEQUENCE FROM N.A.
RC   [illegible]
RX   MEDLINE=[illegible]; PubMed=[illegible];
RA   Akiyama N., [illegible] Sakamoto H., [illegible]
RA   [illegible]
RT   "[illegible] candidate gene [illegible] CAB1 [illegible]
RT   [illegible] ERBB2 amplicon by a modified cDNA selection
RT   [illegible]"
RL   [illegible]
RN   [3]
RP   [illegible]
RX   MEDLINE=[illegible]; PubMed=[illegible];
RA   [illegible]
RT   "[illegible] STAR-related domain."
RL   Nat. [illegible]
CC   -!- FUNCTION: [illegible] AND TRANSFERS [illegible]
CC       [illegible] IN LIVER [illegible] AND BRAIN.
CC   -!- SIMILARITY: [illegible] STAR FAMILY.
CC   --------------------------------------------------------------------------
CC   [illegible]
CC   --------------------------------------------------------------------------
DR   EMBL; [illegible]
DR   EMBL; [illegible]
DR   [illegible]
DR   [illegible]
DR   [illegible]
DR   [illegible]
DR   [illegible]
KW   [illegible]
KW   [illegible]
SQ   SEQUENCE   [illegible] AA;   [illegible] MW;   [illegible] CRC64;

  Query Match        [illegible]
  Best Local Similarity   [illegible]
  Matches [illegible]

Qy   [illegible]
Db   [illegible]

Qy   [illegible]
Db   [illegible]

Qy   [illegible]
Db   [illegible]

Qy   [illegible]
Db   [illegible]
```

```
RESULT  10
ML64_MOUSE
ID   ML64_MOUSE     STANDARD;      PRT;   446 AA.
AC   Q61542;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   MLN 64 PROTEIN (ES 64 PROTEIN).
GN   MLN64 OR ES64.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96039245; PubMed=7490069;
RA   Tomasetto C., Regnier C.H., Moog-Lutz C., Mattei M.-G., Chenard M.-P.,
RA   Lidereau R., Basset P., Rio M.-C.;
RT   "Identification of four novel human genes amplified and overexpressed
RT   in breast carcinoma and localized to the q11-q21.3 region of
RT   chromosome 17.";
RL   Genomics 28:367-376(1995).
CC   -!- FUNCTION: BINDS AND TRANSPORTS CHOLESTEROL. PROMOTES
CC       STEROIDOGENESIS IN PLACENTA AND BRAIN (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE STAR FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; X82457; CAA57834.1; -.
DR   InterPro; IPR000799; -.
DR   InterPro; IPR002913; -.
DR   Pfam; PF01852; START; 1.
DR   PRINTS; PR00978; STARPROTEIN.
KW   Lipid-binding; Lipid transport; Transport; Steroidogenesis.
SQ   SEQUENCE   446 AA;  50469 MW;  DBF43[illegible] 4F7E1E2 CRC64;

  Query Match              24.8%;  Score [illegible];  DB 1;  Length 446;
  Best Local Similarity   36.5%;  Pred. No. [illegible]e-21;
  Matches   76;  Conservative   47;  Mismatches   [illegible];  Indels   [illegible];  Gaps   [illegible];

Qy       68 [illegible] 127
             :| ||  |::| :||          ||: :|    |          : |   | :   |.    | | |: :
Db      232 PSAQFRPYIRQGKEATAVVDQIIAQPENWKFI[illegible]EYGDTVYTIEVPFHGKTFILKTFLP 291

Qy      128 QPMERLYEELVERMEAMCEWNPNVKEIKVLQK[illegible]HELAAEAAGNLV[illegible] 187
             | |  :|:|:: : | |  ||  |     ::||::  :|  :::::::  |||  :|  |||.|:|
Db      292 CPAELVYQEVILQPERMVLWNKTVTACQILQ[illegible]VNV [illegible]

Qy      188 RCAKRRGSTCVLAGMATDFGNMPEQKGVIRAE[illegible]MVLHPLAGSPSKTKLTWLLSIDL 24[illegible]
             |  :||    : :|:||   : |      :| |::        ::     | :|          : : :|
Db      352 PRIPRRRDRY[illegible]IATIHCSKPPTHKYVR[illegible]IVLKTANN[illegible] 41[illegible]

Qy      248 KGWLPKSIINQVLSQTQVPFANHLRKRL 27[illegible]
             || ||: :|:| |   |  :|| |::|:
Db      412 KGRLPRYLIHQSLGATMFEFAFHLRQRV 43[illegible]


RESULT  11
YX70_CAEEL
ID   YX70_CAEEL     STANDARD;      PRT;   [illegible] AA.
AC   O17883;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   HYPOTHETICAL 27.1 KDA PROTEIN F52F12.7 IN CHROMOSOME I.
GN   F52F12.7.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   Matthews L.;
RL   Submitted (DEC-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: BELONGS TO THE STAR FAMILY.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; Z83228; CAB05736.1; -.
DR   WormPep; F52F12.7; CE10852.
DR   InterPro; IPR000799; -.
DR   InterPro; IPR002913; -.
DR   Pfam; PF01852; START; 1.
DR   PRINTS; PR00978; STARPROTEIN.
KW   Hypothetical protein.
SQ   SEQUENCE   241 AA;  27124 MW;  [illegible] CRC64;

  Query Match              12.4%;  Score [illegible];  DB 1;  Length 241;
  Best Local Similarity   26.5%;  Pred. No. [illegible]e-07;
  Matches   57;  Conservative   28;  Mismatches   73;  Indels   47;  Gaps   10;

Qy       65 PPIYS[illegible]AY[illegible]KVVP[illegible]KVPRPV 124
             | ::::|:    ||                 :       |:         | |  : |   '  |:: :   :
Db       40 EAIFNDEN  Y.. --- -----  --SL[illegible]DVVS[illegible]KDIPH[illegible]EMVIISI 81

Qy      125 VVDQPMERLYE[illegible]LVERMEAMCEWNPNVK[illegible]HELAAEAAGN [illegible] LV 17[illegible]
              :  |:|  :  :      |||: |||  ::       .:                     .       :.
Db       82 LL[illegible]ALLEWNQN[illegible] [illegible]V.VV 1[illegible]4

Qy      180 GPRDFV[illegible]AKRRGST[illegible]VLAGMATD[illegible]MVLHPLA[illegible] [illegible]
               |:||| |    :: |     :||    :       :       :  |     :||             |  :||
Db      1[illegible] SGKDFVGAK[illegible]KRKVGLGFILASKGVIV[illegible]KVIAH          LH[illegible] 1[illegible]

Qy      2[illegible]     SPGKIKL[illegible]KGWLPKSIINQVL[illegible]V [illegible]
                 :|    |        ::    ||||:|||  |:||    :         :
Db      1[illegible] [illegible]MLA[illegible]KGYLLKMIVNQV [illegible]

RESULT  12
TNP6_ENTFC
ID   TNP6_ENTFC     STANDARD;      PRT;   988 AA.
AC   Q06238;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   TRANSPOSASE FOR TRANSPOSON TN1546.
OS   Enterococcus faecium (Streptococcus faecium).
OG   Plasmid pIP816.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Enterococcaceae;
OC   Enterococcus.
OX   NCBI_TaxID=1352;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BM4147;
RX   MEDLINE=93[illegible]44; PubMed=8380148;
RA   Arthur M., Molinas C., Depardieu F., Courvalin P.;
RT   "Characterization of Tn1546, a Tn3-related transposon conferring
RT   glycopeptide resistance by synthesis of depsipeptide peptidoglycan
RT   precursors in Enterococcus faecium BM4147.";
RL   J. Bacteriol. 175:117-127(1993).
```

```
CC   -!- FUNCTION: [illegible] TRANS[illegible]
CC   -!- SIMILARITY: BELONGS TO THE [illegible]
CC       TRANSPOSASE.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry [illegible]
CC   between the Swiss Institute of Bioinformatics [illegible]
CC   the European Bioinformatics Institute [illegible]
CC   use by non-profit [illegible]
CC   modified and this statement [illegible]
CC   entities requires a license agreement (See http://[illegible]
CC   or send an email to [illegible]
CC   --------------------------------------------------------------------------
DR   EMBL; M[illegible]29[illegible]; AAA[illegible]
DR   PIR; A4[illegible]28; A[illegible]
DR   InterPro; IPR[illegible]
DR   Pfam; PF[illegible]
KW   Transposable element; Transposition; DNA[illegible]
KW   Plasmid.
SQ   SEQUENCE   [illegible] AA;  [illegible] MW;  [illegible]

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches [illegible]

Qy        41 [illegible]PIPSIW[illegible]NQ[illegible]R [illegible]
                |    |
Db       2[illegible]1 [illegible]PEI[illegible] [illegible]

Qy        [illegible]3      LAY[illegible]AM[illegible]A [illegible]

Db       281 [illegible]YI[illegible] [illegible]

Qy       11[illegible] VEREVVVI[illegible]RLY[illegible]VE[illegible]MEAM[illegible]WNENVE[illegible]
                   |
Db       34[illegible]      QAL[illegible]EAE[illegible]KL[illegible]VEKV[illegible]SVI[illegible]WN [illegible]

Qy       17[illegible] V[illegible]DEV[illegible]V[illegible]AER[illegible]VLA[illegible]MA[illegible]NME[illegible]VIRA[illegible]
                 |:
Db       37[illegible] AEPADY [illegible]

Qy       23[illegible] I[illegible]WLLSILLK[illegible]PKS[illegible]NQV[illegible]ANE[illegible]
                |          |
Db       414 E[illegible]QAVE[illegible]NNES[illegible]EKVE[illegible]W[illegible]

[illegible] 13
ID   [illegible]_HUMAN
ID   [illegible]_HUMAN      STANDARD;      PRT;   [illegible] AA.
AC   [illegible]46[illegible]39;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   [illegible]ROPHIN (DYS[illegible]N-[illegible] PROTEIN) [illegible]
GN   [illegible]RN OR DMDL.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; [illegible]
OC   Mammalia; Eutheria; [illegible]
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=[illegible]; PubMed=[illegible];
RA   [illegible] M., [illegible]
RA   Byth B.[illegible], Knight A.E., Kendrick-Jones J., [illegible]
RA   Edwards Y.H., [illegible]
RT   "Primary structure [illegible]"
RL   Nature [illegible] (199[illegible]).
RN   [2]
RP   X-RAY CRYSTALLOGRAPHY [illegible]
RX   MEDLINE=[illegible]; PubMed=[illegible];
RA   Keep N.H., Norwood F.L.M., Moores C.A., [illegible]
RA   Kendrick-Jones J.;
RT   "The 2[illegible] A structure [illegible]
RT   [illegible]
RL   [illegible] M[illegible] (199[illegible]).
CC   -!- FUNCTION: MAY PLAY A ROLE IN ANCHORING THE CYTOSKELETON TO THE
CC       PLASMA MEMBRANE. [illegible]
CC   -!- [illegible]
CC   -!- TISSUE SPECIFICITY: [illegible]
CC   -!- SIMILARITY: [illegible]
CC       [illegible] MAIN [illegible] IN MANY
CC       ACTIN-BINDING [illegible] SPECTRIN [illegible]
CC       ABP-[illegible] (BETA-[illegible]); [illegible]
CC   -!- SIMILARITY: CONTAINS [illegible] WW [illegible] MAIN.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry [illegible]
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   [illegible]
CC   [illegible]
CC   modified and this statement [illegible]
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   [illegible]
CC   --------------------------------------------------------------------------
DR   EMBL; X[illegible]; AAA[illegible]
DR   [illegible]
DR   [illegible]
DR   MIM; [illegible]
DR   InterPro; IPR[illegible]
DR   InterPro; IPR[illegible]
DR   InterPro; IPR[illegible]
DR   InterPro; IPR[illegible]
DR   InterPro; IPR[illegible]
DR   InterPro; IPR[illegible]
DR   Pfam; PF[illegible]
DR   Pfam; PF[illegible] WW.
DR   Pfam; PF[illegible]
DR   Pfam; PF[illegible]
DR   PROSITE; [illegible] WW_MAIN.
DR   PROSITE; [illegible] ACTININ_1[illegible]
DR   PROSITE; [illegible] ACTININ_2[illegible]
DR   PROSITE; [illegible]
DR   PROSITE; [illegible] WW_DOMAIN_1[illegible]
DR   PROSITE; [illegible] WW_DOMAIN_2[illegible]
KW   Structural protein; Actin-binding; [illegible] Cytoskeleton; [illegible]
KW   [illegible]
FT   [illegible]MAIN [illegible] ACTIN-BINDING.
FT   [illegible]MAIN [illegible] [illegible]
FT   REPEAT     [illegible]
FT   REPEAT     [illegible]
FT   REPEAT     [illegible]
FT   REPEAT     [illegible]
FT   REPEAT     [illegible]
FT   REPEAT     [illegible]
FT   REPEAT     [illegible]
FT   REPEAT     [illegible]
FT   REPEAT     [illegible]
FT   REPEAT     [illegible]
FT   REPEAT     [illegible]
FT   REPEAT     [illegible]
FT   REPEAT     [illegible]
FT   REPEAT     [illegible]
FT   REPEAT     [illegible]
FT   REPEAT     [illegible]
FT   REPEAT     [illegible]
FT   REPEAT     [illegible]
FT   REPEAT     [illegible]
FT   [illegible]MAIN    [illegible]   WW.
FT   [illegible]MAIN    [illegible]   [illegible]MAIN.
FT   [illegible]MAIN    [illegible]
SQ   SEQUENCE   [illegible] AA;  [illegible] MW;  [illegible]

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches [illegible]
Qy      12 SSYRHMENMKCLE[illegible]AVMAISQE[illegible]EEA[illegible]
Db     840 SCQRFE[illegible]NI[illegible]HPKIPMA [illegible]FASCTA[illegible]

Qy      61 [illegible]YSL[illegible]ELAYL[illegible]
Db     891 EAVEDRQQHLENELKS[illegible]BAYLETLKTLEEVL[illegible]

Qy     104 NC[illegible]KVMSKVVPEVGKVFF [illegible]EVVVD[illegible]PME[illegible]
Db     951 KILDEILENQKPALHKLAEETKALEKNVHEDV[illegible]KQ--EFDDV[illegible]KWN     KLKVL [illegible]

Qy     159 KICKDTFITHELAA-----EAAGNLV--- [illegible]VSVECA[illegible]
Db    1004  VSKDLHLLEETALTLEAFEADSTVIEKWHD[illegible]LMKQQAAQ[illegible]

Qy     198 --VLAGMATDFGNMPEQKGVIRAEHGPTCMVL[illegible]SPSKTKLIWLLS[illegible]DLKGW     250
Db    1063 VNEIETIESSLKNMKEIETNLRS------ --G[illegible]VA[illegible]-----  -IKIWV[illegible]1099

Qy     251 ---------LPKSIINQV--LSQTQVDFANHLLK[illegible] 275
          1100 GDYQTQIFKISKEIATQKSELSESQEKAAN LKK[illegible] 1124

RESULT  14
YKU4_YEAST
ID   YKU4_YEAST     STANDARD;      PRT;    [illegible] AA.
AC   P36041;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   HYPOTHETICAL 69.8 KDA PROTEIN IN LOS1 [illegible]2 INTERGENIC REGION.
GN   YKL204W.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Pohl T.M., Pohl F.M.;
RL   Submitted (MAR-1994) to the EMBL/GenBank/DDBJ databases.
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement [illegible]
CC   or send an email to license@isb-sib.ch).
     -----------------------------------------------------------------------
     EMBL; Z28204; CAA82049.1; -.
DR   PIR; S38042; S38042.
DR   SGD; S0001687; YKL204W.
KW   Hypothetical protein; ATP-binding.
FT   NP_BIND     440    447       ATP (POTENTIAL).
SQ   SEQUENCE   [illegible] AA;  69762 MW;  [illegible] CRC64;

  Query Match             6.1%;  Score [illegible];  DB 1;  Length 632;
  Best Local Similarity   21.6%;  Pred. No. [illegible];
  Matches   44;  Conservative   41;  Mismatches   77;  Indels   42;  Gaps    8;

Qy      84 QKALGII SNQE[illegible]WKKESQQ[illegible]NGDKVMSKVVP[illegible] [illegible]
Db     182 EPPMSRIKNMQ[illegible]KRPSQQDDEEKKIDQ[illegible]MLEML[illegible]--------[illegible]LVGHSIAL[illegible]

Qy     144 MGEWNPNVKEIKVLQKIGKDTFITHELAAEAA[illegible]VGPEDEVS      VR[illegible]AKRR[illegible]SIC [illegible]
Db     231 FEEWKAEMEELE LKKLSKSKCISNSTA---      IAPKESASHESI[illegible]LRPVIPKGESS [illegible]

Qy     198 VLAGMAI[illegible]GNMPEQ-KGVIRAEHGPTCMV[illegible]
```

```
Db     [illegible]

Qy     257 [illegible] 280
Db     [illegible]NKSATSLPS 35[illegible]

RESULT  15
YKC2_CAEEL
ID   YKC2_CAEEL     STANDARD;      PRT;    [illegible] AA.
AC   P41993;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   HYPOTHETICAL [illegible] KDA PROTEIN B0280.[illegible] IN CHROMOSOME III.
GN   B0280.2.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN BRISTOL N2;
RA   Fulton L., Waterston R.;
RL   Submitted (JUN-1994) to the EMBL/GenBank/DDBJ databases.
CC   -----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement [illegible]
CC   or send an email to license@isb-sib.ch).
CC   -----------------------------------------------------------------------
DR   EMBL; U10438; AAA19081.1; -.
DR   WormPep; B0280.2; CE00734.
DR   InterPro; IPR001960; -.
DR   Pfam; PF00568; WH1; 2.
KW   Hypothetical protein.
SQ   SEQUENCE   [illegible] AA;  115196 MW;  [illegible] CRC64;

  Query Match             6.1%;  Score [illegible];  DB [illegible];  Length [illegible];
  Best Local Similarity   19.5%;  Pred. No. [illegible];
  Matches   58;  Conservative   50;  Mismatches  135;  Indels   54;  Gaps   10;

Qy      40 AISQELNEEAL[illegible]STWINQVERR[illegible]HLAMLKA[illegible] 89
Db     [illegible]

Qy      90 LS       N[illegible]GWKKESQQ[illegible]NGDKVMSKVV            [illegible]VVVD[illegible]EM [illegible]
Db     562 [illegible]IQSAEPEKI[illegible]IVA[illegible]NA[illegible]WRKK[illegible]AAVAL   [illegible]KNS[illegible]

Qy     131 ERLYEE[illegible]VE[illegible]GEWNPNVKE[illegible]KV[illegible]AA[illegible]GNLV[illegible]EDEVSVE [illegible]
Db     620 ERYYR[illegible]VMQ[illegible]WEEVQPE[illegible]LLV[illegible]NEY[illegible]SKE [illegible]

Qy     189 CAKRRGS[illegible]VLAGMAIDFGNMPEQK[illegible]VI[illegible]KIK               238
Db     680 CENFEN[illegible]KAKQEALE[illegible]EN[illegible]
  
Qy     239    LIWLLSH[illegible]DLKGWLPKSIIN     QV[illegible]EANHLKK[illegible]     [illegible]SHASEAR 284
Db     736 ERHVD[illegible]VK[illegible]VQEELYMLVNNPPP[illegible]KKK[illegible]VKEVK [illegible]


Search completed: August 22, 2001, [illegible]
Job time: 413 sec
```

```
                    GenCore version [illegible]
          Copyright (c) 1993 - 2000 [illegible]mpugen Ltd.


OM protein - protein search, using sw model

Run on:         August 22, 2001, 14:27 2[illegible]  Search time [illegible] Seconds
                                          (without alignments)
                                          [illegible]

Title:          US-08-659-254-18
Perfect score:  1470
Sequence:       1 MLLATFKL[illegible]YRHMRNM ..........[illegible]

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.[illegible]

Searched:       425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters:      425[illegible]

Minimum DB seq length: 0
[illegible]mum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       SPTREMBL_16:*
               1:  sp_archea:*
               2:  sp_bacteria:*
               3:  sp_fungi:*
               4:  sp_human:*
               5:  sp_invertebrate:*
               6:  sp_mammal:*
               7:  sp_mhc:*
               8:  sp_organelle:*
               9:  sp_phage:*
               10:  sp_plant:*
               11:  sp_rodent:*
               12:  sp_unclassified:*
               13:  sp_vertebrate:*
               14:  sp_virus:*

    Pred. No. is the number of results predicted by chance to have a
    score greater than or equal to the score of the result being printed,
    and is derived by analysis of the total score distribution.

                              SUMMARIES
              *
[illegible]t          Query
[illegible]o.  Score  Match Length DB  ID                     Description
------------------------------------------------------------------------------
  1    1295   88.1    285  6  Q9GMD0                 Q9gmd0 ovis aries
  2    1069   72.7    281 13  Q9DG09                 Q9dg09 gallus gal
  3    1045   71.1    289 13  Q9DG08                 Q9dg08 xenopus lae
  4   990.5   67.4    287 13  Q9DFB4                 Q9dfb4 oncorhyncho
  5   985.5   67.0    287 13  [illegible]            [illegible]
  6     969   65.9    285 13  Q9DC10                 Q9dc10 brachydanio
  7   [illegible] 2[illegible] [illegible] 13 [illegible] [illegible]
  8   229.5   15.6    580  5  Q9W145                 Q9w145 drosophila
  9     189   12.[illegible] 219 11 Q9EPQ7           Q9epq7 mus musculu
 10     159   10.8    180 11  Q9EPQ[illegible]       Q9epq[illegible] mus musculu
 11     130    8.8    478  5  Q198[illegible]9       Q198[illegible]9 caenorhabdi
 12   119.5    8.1    106  4  Q9NSY2                 Q9nsy2 homo sapien
 13   109.5    7.4    995  4  Q9Y3M8                 Q9y3m8 homo sapien
 14     109    7.4   1083  4  O14868                 O14868 homo sapien
 15   108.5    7.4    598 11  [illegible]            [illegible]
 16     108    7.3   1091  4  O431[illegible]9       O431[illegible]9 homo sapien
 17   107.5    7.[illegible] 219 1 [illegible]       [illegible]
 18   107.5    7.3   1281  2  Q912[illegible]1       Q912[illegible]1 pseudomonas
 19   103.5    7.0   1047  5  Q9W3V4                 [illegible]
 20   101.5    6.9    239  1  O30224                 O30224 archaeoglob
 21     101    6.9   1083 11  Q63744                 Q63744 rattus norv
 22     100    6.8    399  1  O58288                 O58288 pyrococcus
 23    99.5    6.8    624 11  Q9EQ[illegible]9       Q9eq[illegible]9 mus musculu
 24      99    6.7    359  4  Q9Y3[illegible]5       Q9y3[illegible]5 homo sapien
 25    98.5    6.7    291 11  Q9[illegible]          Q9[illegible] mus musculu
 26    98.5    6.7    663 10  Q9FH45                 Q9fh45 arabidopsis
 27      97    6.6    970  5  Q9N8Y9                 Q9n8y9 trypanosoma
 28    96.5    6.6    679 10  O49417                 O49417 arabidopsis
 29      96    6.5    256  5  Q21130                 Q21130 caenorhabdi
 30      96    6.5    419  1  Q9UY16                 Q9uy16 pyrococcus
 31      95    6.5   1092 11  Q9R[illegible]         Q9r[illegible] mus musculu
 32    94.5    6.4    624  4  Q9Y5P4                 Q9y5p4 homo sapien
 33      94    6.4     22  5  [illegible]            [illegible] caenorhabdi
 34      94    6.4    469  2  O67452                 O67452 aquifex aeo
 35      94    6.4    753  1  Q9YFF2                 Q9yff2 aeropyrum p
 36    93.5    6.4    598  4  Q9H2S8                 Q9h2s8 homo sapien
 37    93.5    6.4    598  6  Q9GK16                 Q9gk16 bos taurus
 38    93.5    6.4    624  6  Q9GK17                 Q9gk17 bos taurus
 39      93    6.3    354  1  O59485                 O59485 pyrococcus
 40      91    6.2   1520 10  Q39[illegible]6        Q39[illegible]6 arabidopsis
 41      90    6.[illegible] 410 2 Q9KH26            Q9kh26 coxiella bu
 42      90    6.1   1188  2  Q9KA07                 Q9ka07 bacillus ha
 43      90    6.1   3309  5  Q9GUP2                 Q9gup2 caenorhabdi
 44      90    6.1   3419 11  O55147                 O55147 rattus norv
 45      89    6.1    607 10  Q9S143                 Q9s143 arabidopsis


                              ALIGNMENTS


RESULT 1
Q9GMD0
ID   Q9GMD0      PRELIMINARY;      PRT;   285 AA.
AC   Q9GMD0;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN.
GN   STAR.
OS   Ovis aries (Sheep).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Caprinae; Ovis.
OX   NCBI_TaxID 9940;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   West L.A., Horvat R.D., Roess D.A., [illegible] B.G., [illegible] J.L.,
RA   Niswender G.D.;
RT   "Steroidogenic Acute Regulatory Protein and Peripheral-type
RT   Benzodiazepine Receptor Associate at the Mitochondrial Membrane.";
RL   Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF290202; AAG02464.1; -.
SQ   SEQUENCE   285 AA;  31895 MW;  AB[illegible]A[illegible]F295[illegible] CRC64;

  Query Match             88.1%;  Score [illegible];  DB 6;  Length 285;
  Best Local Similarity   87.4%;  Pred. No. [illegible];
  Matches  249;  Conservative   [illegible]9;  Mismatches   17;  Indels    0;  Gaps    0;

Qy          1 MLLATFKL[illegible]YRHMRNMK[illegible]KVEEESS[illegible]
              [illegible]
Db          1 MLLATFKL[illegible]YRHVRSMKGLRQ[illegible]
 
Qy         61 SRLEETLYS[illegible]QELAYLQQGEEAMQKAL[illegible]KVMSKVVPDVGKVF 120
              [illegible]
Db         61 SQLEDSLYS[illegible]QELAYIQQGEEAMQRAL[illegible]KVIPDVGKVF 120

Qy        121 KLEVVVL[illegible] 180
              [illegible]
Db        121 KLEVVV[illegible] 180

Qy        181 [illegible]SKIK[illegible] 240
```

```
            [illegible]
Db    181 [illegible]DFVSV[illegible]

Qy    241 WLESI[illegible]K[illegible]W[illegible]K[illegible]NQV[illegible]V[illegible]ANH[illegible]RK[illegible]
            [illegible]
Db    241 WLESIDLK[illegible]W[illegible]KT[illegible]NQV[illegible]V[illegible]ANH[illegible]


RESULT   [illegible]
Q9DG09
ID   Q9DG09      PRELIMINARY;      PRT;   28[illegible] AA.
AC   Q9DG09;
DT   01 MAR 2001 (TrEMBLrel. [illegible], Created)
DT   01 MAR 2001 (TrEMBLrel. [illegible], Last sequence update)
DT   01 MAR 2001 (TrEMBLrel. [illegible], Last annotation update)
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN
GN   STAR.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; [illegible]
OC   Archosauria; Aves; Neognathae; [illegible]
OC   Gallus.
OX   NCBI_TaxID=[illegible]
RN   [1]
RP   SEQUENCE FROM N.A.
RX   PubMed=1106[illegible]
RA   Bauer M.[illegible], Bridgham [illegible]
RT   "Conservation of steroidogenic acute [illegible]
RT   structure and expression in vertebrates."
RL   Mol. Cell. Endocrinol. [illegible]
DR   EMBL; AF2204[illegible]; AAG2[illegible]
SQ   SEQUENCE   28[illegible] AA;  [illegible] MW;  [illegible]

  Query Match             [illegible]
  Best Local Similarity   [illegible];  Pred. No. [illegible]
  Matches  20[illegible];  Conservative   4[illegible];  Mismatches  [illegible]

Qy      1 MLLATFKL[illegible]A[illegible]YR[illegible]NMK[illegible]RQLAVMA[illegible]NRPAL[illegible]
            [illegible]
Db      1 MLPATFKL[illegible]AA[illegible]YQR[illegible]NV[illegible]LRRQAAVA[illegible]

Qy     61 SRLEEILYS[illegible]AYI[illegible]AM[illegible]KA[illegible]N[illegible]W[illegible]
            [illegible]
Db     61 SRLEEKPE[illegible]M[illegible]SYI[illegible]A[illegible]K[illegible]

Qy    121 R[illegible]VVVI[illegible]MEL[illegible]YEE[illegible]EMPAM[illegible]WN[illegible]NV[illegible]
            [illegible]
Db    121 R[illegible]VVVI[illegible]

Qy    181 PRGEVSVR[illegible]AK[illegible]VLA[illegible]MAIL[illegible]NMEL[illegible]
            [illegible]
Db    181 PRGEVSVR[illegible]SIL[illegible]VLA[illegible]MSIL[illegible]AMLR[illegible]RAIN[illegible]

Qy    241 WLESIDLK[illegible]W[illegible]EN[illegible]V[illegible]QVLLANH[illegible]
            [illegible]
Db    241 WLESIDLK[illegible]W[illegible]IN[illegible]V[illegible]QLVLLAH[illegible]EM[illegible]AR[illegible]


RESULT   3
Q9DG08
ID   Q9DG08      PRELIMINARY;      PRT;   [illegible] AA.
AC   Q9DG08;
DT   01 MAR 2001 (TrEMBLrel. [illegible], Created)
DT   01 MAR 2001 (TrEMBLrel. [illegible], Last sequence update)
DT   01 MAR 2001 (TrEMBLrel. [illegible], Last annotation update)
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN [illegible] FRAGMENT.
GN   STAR.
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; [illegible]
OC   Amphibia; Batrachia; Anura; Mesobatrachia; [illegible]
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=83[illegible];
RN   [1]
```

```
RP   SEQUENCE FROM N.A.
RC   [illegible] OVARY.
RX   [illegible]
RA   Bauer M.[illegible], Bridgham [illegible]
RT   "Conservation of steroidogenic acute regulatory (STAR) protein
RT   structure and expression in vertebrates."
RL   Mol. Cell. Endocrinol. [illegible]
DR   EMBL; AF[illegible]; AAG[illegible]
DR   [illegible]
DR   [illegible]
SQ   SEQUENCE   [illegible] AA;  [illegible] MW;  [illegible]

  Query Match             [illegible]
  Best Local Similarity   [illegible];  Pred. No. [illegible]
  Matches  [illegible];  Conservative   4[illegible];  Mismatches  [illegible]

Qy      1 M[illegible]A[illegible]K[illegible]A[illegible]NMK[illegible]LAVMA[illegible]WIN[illegible]VRRR[illegible]
            [illegible]
Db      1 [illegible]A[illegible]NMT[illegible]AAVA[illegible]WIN[illegible]RR[illegible]

Qy     61 [illegible]AM[illegible]KA[illegible]N[illegible]W[illegible]N[illegible]KVMSKVVVLV[illegible]KV[illegible]
            [illegible]
Db     61 [illegible]NLVEM[illegible]A[illegible]N[illegible]W[illegible]VMEN[illegible]KV[illegible]KV[illegible]

Qy    121 [illegible]VVV[illegible]EME[illegible]MPAM[illegible]WNENVRE[illegible]RRAAL[illegible]AN[illegible]
            [illegible]
Db    121 [illegible]EAV[illegible]NMERM[illegible]WNENVREVE[illegible]RRRAAL[illegible]N[illegible]V[illegible]

Qy    181 [illegible]EVSVR[illegible]AK[illegible]MA[illegible]NMEL[illegible]MVLHLLA[illegible]
            [illegible]
Db    181 [illegible]AR[illegible]V[illegible]AIMS[illegible]GM[illegible]MVLRLLA[illegible]VGKIKL[illegible]

Qy    241 WLESI[illegible]K[illegible]WLEK[illegible]N[illegible]V[illegible]QVLLANHLRKRLLH[illegible]ARL 28[illegible]
            [illegible]
Db    24[illegible] WLESI[illegible]K[illegible]WLEK[illegible]NQV[illegible]QLVLLAKHLRSRMA[illegible]AL[illegible] 28[illegible]


RESULT   4
Q[illegible]DE[illegible]4
ID   [illegible]      PRELIMINARY;      PRT;   28[illegible] AA.
AC   [illegible];
DT   01 MAR 2[illegible] (TrEMBLrel. [illegible], Created)
DT   [illegible] MAR 2[illegible] (TrEMBLrel. [illegible], Last sequence update)
DT   [illegible] MAR 2[illegible] (TrEMBLrel. [illegible], Last annotation update)
DE   STEROIDOGENIC ACUTE REGULATORY PROTEIN
GN   STAR.
OS   [illegible]
OC   Eukaryota; Metazoa; [illegible]
OC   [illegible]
OC   [illegible]
OX   NCBI_TaxID=[illegible]
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=HEAD [illegible] AND KIDNEY
RA   [illegible]
RT   "M[illegible] acute regulatory (STAR) protein
RT   mRNA from rainbow trout."
RL   Submitted [illegible] to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB[illegible]; BAB[illegible]
SQ   SEQUENCE   [illegible] AA;  [illegible] MW;  [illegible]

  Query Match             [illegible]
  Best Local Similarity   [illegible];  Pred. No. [illegible]
  Matches  [illegible];  Conservative   [illegible];  Mismatches  [illegible];  Indels  [illegible];  Gaps  [illegible]

Qy      1 MLLA[illegible]NMK[illegible]LAVMA[illegible]WIN[illegible]VRRR[illegible]
            [illegible]
Db      1 MLPA[illegible]NM[illegible]RNAMVA[illegible]N[illegible]HVRRR[illegible]

Qy     61 [illegible]AM[illegible]N[illegible]W[illegible]N[illegible]KVMSKVVLV[illegible]KV[illegible]
            [illegible]
```

```
Db       59 SRIEEEQGYNEAFVSYVKQGEEALQKSISIL[illegible]WIIEIIAANCEV[illegible] 11[illegible]

Qy      120 FELEVVVDQPMEELYEELVERMEAMGEWNENVKE[illegible]IKVLQKIGKD[illegible]HHLAAHAADNLV 17[illegible]
            |:|||::||    :  || |||    ||  ||:||:|:|   :  ::||||::|  :|||::        |::|
Db      119 FKLEVLLDQRSDNLYVELVGNMEQMGDWNENVKEVK[illegible]LKIGQEIMVILLVC[illegible]INVV 17[illegible]

Qy      180 GPRDFVSVRCAKRRGSTCVLACMATDFGNMPE[illegible]KGVIRABRGPTCMVLHPLAKTPCKIKL 2[illegible]
            ||||||||||||||||  |||||:|      ||  :|  :||  :||||:::: | |  |:||:
Db      179 GPRDFVSVRCAKRRGSTCPLACMSTQHPTMPE[illegible]VVRAENGPTCIVMRPCALIENKIKF [illegible]

Qy      240 TWLLSIDLKGWLPKSIINQVLSQTQVDFANHL[illegible]KRLLSHPAC 28[illegible]
            ||||||||||:||:|||:||||||||||||||  :      :     .
Db      239 TWLLSIDLKGWIPKTIINKVLSQTQVDFANHL[illegible]EMALNGVS 28[illegible]


RESULT    5
Q9DE06
ID    Q9DE06         PRELIMINARY;       PRT;   2[illegible] AA.
AC    Q9DE06;
DT    01-MAR-2001 (TrEMBLrel. 16, Created)
DT    01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT    01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
      STEROIDOGENIC ACUTE REGULATORY PROTEIN
      STAR.
      Salvelinus fontinalis (Brook trout) [illegible]
OC    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC    Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC    Protacanthopterygii; Salmoniformes; Salmonidae; Salvelinus.
OX    NCBI_TaxID=8038;
RN    [1]
RP    SEQUENCE FROM N.A.
RC    TISSUE=BRAIN, TESTIS, HEAD, KIDNEY, SKIN, AND LIVER;
RA    Goetz F.W.;
RT    "Characterization of the brook trout StAR and MLN64 homologs.";
RL    Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.
DR    EMBL; AF232215; AAG39689.1; -.
SQ    SEQUENCE   287 AA;  31994 MW;  [illegible] CRC64;

  Query Match             67.0%;  Score 9[illegible];  DB 13;  Length 287;
  Best Local Similarity   65.4%;  Pred. No. 1.6e-[illegible];
  Matches  185;  Conservative   46;  Mismatches   49;  Indels    3;  Gaps    2;

Qy        1 MLLATFKLCAGSSYRHMENMEGLEQQAVMA[illegible]NSEALGCL[illegible]
            || |||||||| |||| ||| |||: |::         |:|:::| |||| |
Db        1 MLPATFKLCACISYRHIRNMIGLEKNAPVA[illegible]MLACPNPSSWIRRVEEKRGLLS[illegible]

Qy       61 SRLEFTL[illegible]SLQSLAYLQQGEEAMQKAL[illegible]
            ||:||    |::  |::|::|||||:||:          |   |   :|| :|   ||
Db       59 SRIEEEQCYNEAFVSYVKQGEEALQKS[illegible]IIAANCEKVLCKVLEIVCKV 11[illegible]

Qy      120 FELEVVVDQPMEELYEELVERMEAMGEWNPNVK[illegible]IKVLQKIGKD[illegible]HHLAAHAADNLV 17[illegible]
            |:|||::||    :  || |||    ||  ||:||:|:|   :  ::||||::|  :|||::        |::|
Db      119 FKLEVLLDQRSDNLYGELVGNMEQMGDWNPNVKEVKILQKIGQEIMVIHEVS[illegible]INVV 1[illegible]

Qy      180 GPRDFVSVRCAKRRGSTCVLACMATDFGNMP[illegible]KGVIRABRGPTCMVLHPLAKTPCKIKL 2[illegible]
            ||||||||||||||||||||  ||||:|      ||||  ||:|||:||||::  | |      :||:|
Db      179 GPRDFVSVRCAKRRGSTCPLACMSTQHPTMP[illegible]VVRAENGPTCIVMRPCALIENKIKF [illegible]

Qy      240 TWLLSIDLKGWLPKSIINQVLSQTQVDFANHL[illegible]KRLLSHLAC 2[illegible]
            ||||||||||||:||:|||:||||||||||||  ||   |:   :   |:
Db      239 TWLLSIDLKGWIPKTIINKVLSQTQVDFANHLHEMALNGVGK 28[illegible]


RESULT    6
Q9DG10
ID    Q9DG10         PRELIMINARY;       PRT;   2[illegible]5 AA.
AC    Q9DG10;
DT    01-MAR-2001 (TrEMBLrel. 16, Created)
DT    01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT    01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE    STEROIDOGENIC ACUTE REGULATORY PROTEIN.
GN    STAR.
```

```
OS    Brachydanio rerio (Zebrafish) (Zebra danio).
OC    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC    Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC    Cypriniformes; Cyprinidae; Rasborinae; Danio.
OX    NCBI_TaxID=7955;
RN    [1]
RP    SEQUENCE FROM N.A.
RX    PubMed 11064158;
RA    Bauer M.P., Bridgham J.T., Langenau D.M., Johnson A.L., Goetz F.W.;
RT    "Conservation of steroidogenic acute regulatory (StAR) protein
RT    structure and expression in vertebrates.";
RL    Mol. Cell. Endocrinol. 168:119-125(2000).
DR    EMBL; AF220435; AAG28593.1; -.
SQ    SEQUENCE   285 AA;  31759 MW;  [illegible] CRC64;

  Query Match             65.9%;  Score [illegible];  DB 1[illegible];  Length 285;
  Best Local Similarity   64.6%;  Pred. No. [illegible];
  Matches  179;  Conservative   46;  Mismatches   50;  Indels    2;  Gaps    1;

Qy        1 MLLATFKLCAG[illegible]SYRHMENMEGLEQQAVM[illegible]IWINQVRRESGL[illegible] 6[illegible]
            || ||||||||  |||||||| |||:  ::  |  |:  :        :  | :  :|:|:
Db        1 MLPATFKLCA[illegible]SYRHMRNMIGLEKNAMIA[illegible]HHELNR[illegible]LSGPSA[illegible]IWINHIRRSSLLS 58

Qy       61 SRLEL[illegible]AELLAMQEA[illegible]EVMSKVVDPVCRVF 12[illegible]
            | :  |  |:: :  |:|||:||:|||::  |           |   |:|::|  |  :|:|
Db       59 SEIAEEIYGEA[illegible]YVQQGLAIQKSIGIII[illegible]EVMSKVIPIVCKVF 11[illegible]

Qy      121 KLEVVVD[illegible]MEELYEELVERMEAMGEWNPN[illegible]KVLQKIGKD[illegible]HHLAAHAAGNLV 18[illegible]
            :|||  ::|       :|:||||:  ||  ||  |  ||    |::|:|    ||||:  ||| ::|:  ||:|
Db      119 KLEVTLEQ[illegible]GDLYDELVDNMEQMGEWNPN[illegible]KLVK[illegible]LKIGQ[illegible]IHEEVALHPGNVVG 17[illegible]

Qy      181 PRDFVSVRCAKRRGSTCVLACMATDFGNMP[illegible]KGVIRAEHGPTCMVLHPLAKTPSKIKL 240
            |||||:|   ||||||||  |||:|                    :  :  :|  |  :  :  |   |:||||
Db      179 PRDFVNVRRAKRRGST[illegible]PLACMST[illegible]LKGVVRAENGPTCIVMRPCAMIENKIKF 238

Qy      241 WLLSIDLKGWLPKSIINQVLSQTQVDFANHL[illegible]KRLESC 277
            ||||:||||:|||::|||:|||||||||    |::
Db      239 WLLSIDLKGWIPKIVINRVLSQTQVDFVN[illegible]HMAS 275


RESULT    7
Q9DFS4
ID    Q9DFS4         PRELIMINARY;       PRT;   2[illegible] AA.
AC    Q9DFS4;
DT    01-MAR-2001 (TrEMBLrel. 16, Created)
DT    01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT    01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE    MLN64-LIKE PROTEIN (FRAGMENT).
GN    MLN64.
OS    Brachydanio rerio (Zebrafish) (Zebra danio).
OC    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC    Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC    Cypriniformes; Cyprinidae; Rasborinae; Danio.
OX    NCBI_TaxID=7955;
RN    [1]
RP    SEQUENCE FROM N.A.
RX    PubMed=11064158;
RA    Bauer M.P., Bridgham J.T., Langenau D.M., Johnson A.L., Goetz F.W.;
RT    "Conservation of steroidogenic acute regulatory (StAR) protein
RT    structure and expression in vertebrates.";
RL    Mol. Cell. Endocrinol. 168:119-125(2000).
DR    EMBL; AF258786; AAG28603.1;
FT    NON_TER       1      1
SQ    SEQUENCE   2[illegible] AA;  27345 MW;  [illegible] CRC64;

  Query Match             26.5%;  Score [illegible];  DB [illegible];  Length 2[illegible];
  Best Local Similarity   36.6%;  Pred. No. [illegible];
  Matches   8[illegible];  Conservative   51;  Mismatches   [illegible];  Indels   [illegible];  Gaps    1;

Qy       53 RRRAGLL[illegible]KEAL[illegible]WRRF[illegible]LLNCLRVMSKV 112
            | |  '  '       ::|| |:::|'          '|: :    | :    : ||    :
```

```
Db      [illegible] RVRDD[illegible] GREAV[illegible]KAEVR[illegible]REAMAVV[illegible]ENW[illegible]

Qy      11[illegible] VPDVGKV[illegible]VVV [illegible]PM[illegible]RLYE[illegible]AM[illegible]EWN[illegible]
                  [illegible]
Db      69 [illegible]PHGK[illegible]AL[illegible]VYQ[illegible]V[illegible]QEEKMV[illegible]WN[illegible]

Qy      173 EAAGNLV[illegible]PV[illegible]AKERG[illegible]V[illegible]A[illegible]MA[illegible]NM[illegible]
                  [illegible]
Db      129 [illegible]AAGGVV[illegible]VN[illegible]VE[illegible]MA[illegible]N[illegible]

Qy      233 [illegible]K[illegible]KE[illegible]W[illegible]W[illegible]N[illegible]ANH[illegible]
                  [illegible]
Db      189 N[illegible]VC[illegible]W[illegible]N[illegible]YI[illegible]AAIM[illegible]


RESULT  [illegible]
[illegible]45
ID   Q9W145     PRELIMINARY;     [illegible]     AA.
AC   Q9W14[illegible];
DT   01 MAY 2000 (TrEMBLrel. [illegible], Created)
DT   01 MAY 2000 (TrEMBLrel. [illegible], Last sequence update)
DT   01 MAR 2001 (TrEMBLrel. [illegible], Last annotation update)
DE   CG35[illegible] PROTEIN.
GN   CG35[illegible].
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; [illegible]
OC   Pterygota; Neoptera; [illegible]
OC   Ephydroidea; Drosophilidae; [illegible]
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RX   MEDLINE=20196006; PubMed=[illegible];
RA   Adams M.D., Celniker [illegible]
RA   Amanatides [illegible]
RA   George R.A., [illegible]
RA   Sutton G.G., Wortman [illegible]
RA   Brandon R.C., Rogers [illegible]
RA   Wan K.H., Doyle [illegible], Baxter [illegible]
RA   Abril [illegible]
RA   Ballew R.M., [illegible]
RA   Beeson K.Y., [illegible]
RA   Borkova [illegible]
RA   Burtis K.C., [illegible]
RA   Cherry J.M., [illegible]
RA   de Pablos B., [illegible]
RA   Dodson K., [illegible]
RA   Durbin K.J., [illegible]
RA   Fosler [illegible]
RA   Glodek A., [illegible]
RA   Harris N.L., [illegible]
RA   Hostin [illegible]
RA   Jalali M., [illegible]
RA   Kimmel B.E., [illegible]
RA   Lasko P., Lei Y., [illegible]
RA   [illegible] X., Mattei B., [illegible]
RA   Merkulov [illegible]
RA   Mount [illegible]
RA   Nelson [illegible]
RA   Palazzolo M., [illegible]
RA   Reinert K., [illegible]
RA   Shue B.[illegible]
RA   Spier E., [illegible]
RA   Svirskas R., [illegible]
RA   Wang Z.[illegible], Wassarman [illegible]
RA   Williams S.M., [illegible]
RA   Ye J., Yeh R.[illegible]
RA   Zheng X.H., Zhong F.N., [illegible]
RA   Gibbs R.A., Myers E.W., [illegible]
RT   "The genome sequence of Drosophila [illegible]";
RL   Science 287:2185-[illegible](2000).
DR   EMBL; AE00346[illegible]; AA[illegible].
DR   HSSP; Q[illegible]; [illegible].
```

```
DR   [illegible]
DR   [illegible]
DR   [illegible]
DR   [illegible]
DR   [illegible]
SQ   SEQUENCE [illegible] AA; [illegible] MW; [illegible]

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches [illegible]; Conservative [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible]

Qy      [illegible]AMQKA[illegible]NQ[illegible]WKE[illegible]

Db      [illegible]MD[illegible]EKAYE[illegible]WK[illegible]

Qy      [illegible]VVV[illegible]MF[illegible]AM[illegible]WNENVFE[illegible]KV[illegible]

Db      [illegible]KVNQ[illegible]KYLAKA[illegible]EKWN[illegible]

Qy      [illegible]VLAGMA[illegible]N[illegible]

Db      [illegible]

Qy      [illegible]M[illegible]V[illegible]

Db      [illegible]

Qy      [illegible]

Db      [illegible]

Qy      [illegible]NL[illegible]VE[illegible]ANH[illegible]

Db      [illegible]


RESULT  [illegible]
[illegible]
ID   [illegible]     PRELIMINARY;     [illegible]
AC   [illegible]
DT   [illegible] MAR [illegible] (TrEMBLrel. [illegible], Created)
DT   [illegible] MAR [illegible] (TrEMBLrel. [illegible], Last sequence update)
DT   [illegible] MAR [illegible] (TrEMBLrel. [illegible], Last annotation update)
DE   [illegible] (FRAGMENT).
OS   [illegible]
OC   [illegible]
OC   [illegible]
OX   NCBI_TaxID=[illegible];
RN   [illegible]
RP   SEQUENCE FROM N.A.
RC   STRAIN=[illegible]
RA   [illegible]
RA   [illegible]
RT   "[illegible] development [illegible]
RT   [illegible]";
RL   [illegible]
DR   EMBL; A[illegible]
DR   [illegible]
SQ   SEQUENCE [illegible] AA; [illegible] MW; [illegible]

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches [illegible]

Qy      [illegible]AMQKA[illegible]

Db      [illegible]

Qy      [illegible]

Db      [illegible]
```

```
Qy     193 -PCSTCVLACMATDPCNMPEQKCVIPAEHCPTCMV [illegible] PLACSPSKIEI[illegible] 251
               |          :      | : | :|   : |         || |  :|: |          |: :|
Db     131 YEDCIISSNATHVEHPL[illegible]PPKPGPVRCFNHP[illegible] 19

Qy     252 PKSIINQVLSQIQVDPANHLRK 273
            |:|:::     ::  :|   :|:|
Db     191 PQSVVDSFFPRSMAEFYPNLQK 212


RESULT  10
Q9EPQ6
ID   Q9EPQ6      PRELIMINARY;      PRT;   [illegible] AA.
AC   Q9EPQ6;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   STAR-RELATED PROTEIN F3-15E/F3 16E (FRAGMENT).
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
[illegible]   [1]
[illegible]   SEQUENCE FROM N.A.
[illegible]   STRAIN=SWISS WEBSTER/NIH;
RA   Wines M.E., Lee L., Katari M.S., Zhang [illegible], DeRossi [illegible], Shi Y.,
RA   Perkins S., Feldman M., McCombie W.R. [illegible] B.C.;
RT   "Identification of mesoderm development (mesd) candidate genes by
RT   comparative mapping and genome sequence analysis.";
RL   Submitted (AUG 2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY007809, AAG41057.1; -.
FT   NON_TER       1       1
SQ   SEQUENCE   180 AA;  20099 MW;  AAA576[illegible] CRC64;

  Query Match              10.8%;  Score 15[illegible];  DB 11;  Length 18[illegible];
  Best Local Similarity    25.2%;  Pred. No. [illegible];
  Matches   40;  Conservative   33;  Mismatches   84;  Indels   [illegible];  Gaps    2;

Qy     117 CKVFPLEVVVD[illegible]PMERLYEELVERMEAM[illegible] 175
             | ::| | ::      | ::: :         : |       |    :::| |        ::       ||
Db      15 GNLYFC[illegible]FVW[illegible]IEPVASGL[illegible]AA 74

Qy     176 CNLVCPPDFVSVRCAKE RCSTCVLACMATD[illegible]KCVIPAEHC[illegible]MVLHPLAGSP 234
              |: |||| :    |:      |        :      | : | :|  : |         .| | :
Db      75 MKLISPRDFVDLVLVKKYEDCTISSNA[illegible]RVE[illegible]PLKPGPVRCFNHH[illegible]PLPGDP [illegible]

Qy     235 SKIKLIWLLSIDLKGWLPKSIINQVLSQIQV[illegible]ANHLRK 273
            :|| |        || |:||:|:::     ::  :|   :|:|
Db     135 NKTNLVTFFQTDLSCYLPQSVVDSFFPRSMAE-YPNLQK 173

[illegible]LT  11
Q19819
ID   Q19819      PRELIMINARY;      PRT;   47[illegible] AA.
AC   Q19819;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   COSMID F26F4.
GN   F26F4.4.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RX   MEDLINE=94150718; PubMed=7906398;
RA   Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA   Bonfield J., Burton J., Connell M., Copsey T., Cooper J., [illegible] A.,
RA   Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA   Gardner A., Green P., Hawkins T., Hillier L., Jier M., [illegible] L.,
```

```
RA   Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA   Lightning J., Lloyd C., Mcmurray A., Mortimore B., O'Callaghan M.,
RA   Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,
RA   Smaldon N., Smith A., Sonnhammer E., Staden R., Sulston J.,
RA   Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA   Watson A., Weinstock L., Wilkinson Sproat J., Wohldman P.;
RT   "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT   elegans.";
RL   Nature 368:32-38(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   Fulton L.;
RL   Submitted (NOV 1994) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   Waterston R.;
RL   Submitted (MAR 1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U12964, AAA91219.1; -.
DR   InterPro; IPR002913; -.
DR   Pfam; PF01852; START; 1.
DR   SMART; SM00234; START; 1.
SQ   SEQUENCE   478 AA;  54924 MW;  7F[illegible]EE40[illegible] CRC64;

  Query Match              8.8%;  Score 1[illegible];  DB 5;  Length 478;
  Best Local Similarity    20.2%;  Pred. No. [illegible];
  Matches   62;  Conservative   60;  Mismatches  145;  Indels   40;  Gaps   12;

Qy       1 MLLATFKL[illegible]A[illegible]SYKHMKNMKGLKQQAVMA[illegible]TWINQVFSFSSL[illegible] 60
             ::: :| |       |          |     ::           | | |       :       ||:
Db     188 L[illegible]WSHEY   -   LMPF[illegible]ESNNRH 239

Qy      61 SRLEETLYS[illegible]LAYLQQGEEAMQKA[illegible]KVMQK   VVP[illegible]VGK 118
              |        :                :   : |   :  :      ::   : : :  |:     |    :
Db     240 KRGKKQQNK[illegible]LAFTAVI[illegible]KE 299

Qy     119 VFKLEVVV[illegible]QPMEKLYEELVERMEAM[illegible]     NENVK         E[illegible]KVLQKI[illegible]K 16[illegible]
             :|   |: :  : : ::: |||: |    :: :         | |               | :|: :   :
Db     300 LF- EPILRECLDQV EELM RDSRIGGW[illegible]KMIAP[illegible] [illegible]

Qy     163 [illegible]VIRA 218
              |  :  :  ::|  |     :  |||: :|            |  .           : :     .   : :|
Db     356 AIDLYYSVSAIAMRGYISSRDFLDL[illegible]YEVS[illegible]SNECPNNCNPKIVK[illegible] 41[illegible]

Qy     219 EHCPTCMVLHPLACSPSKIKLIWLL[illegible]PANH[illegible]KKLESH 27[illegible]
              : |:   ::   |          |     ||:    |      . ::   :          |: :: |'
Db     41[illegible] HNFTS   MIKILKLEACVTYFEWLMKT[illegible]KKMNEF 4[illegible]

Qy     279 PASEARC 285
                :   |
Db     470 AETHYHC 476


RESULT  12
Q9NSY2
ID   Q9NSY2      PRELIMINARY;      PRT;   10[illegible] AA.
AC   Q9NSY2;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   HYPOTHETICAL 12.[illegible] KDA PROTEIN.
GN   DKFZP434G2416.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=TESTIS;
RA   Koehrer K., Beyer A., Mewes H.W., [illegible] J., Wiemann S.;
RL   Submitted (JAN 2000) to the EMBL/GenBank/DDBJ databases.
```

```
DR   EMBL; AL[illegible]; CAB[illegible]
KW   Hypothetical protein
SQ   SEQUENCE [illegible] AA; [illegible] MW; [illegible]

  Query Match [illegible]
  Best Local Similarity [illegible] 28.[illegible] Pred. No. [illegible]
  Matches [illegible] 18; Conservative [illegible] Mismatches [illegible] Gaps [illegible]

Qy   178 LVGPRDEV[illegible]AKE[illegible]VLAGMAIDE[illegible]NMF[illegible]
           |||
Db     3 LLSPRDEV[illegible]VKE[illegible]NA[illegible]VEDE[illegible]

Qy   237 EKEIWE[illegible]WE[illegible]NLVE[illegible]EANGE[illegible]
Db    63 [illegible]NEVIEE[illegible]YE[illegible]NVV[illegible]AN[illegible]

RESULT 13
Q9Y3M8
ID   Q9Y3M8     PRELIMINARY; [illegible] AA.
AC   Q9Y3M8;
DT   01-NOV-1999 (TrEMBLrel. [illegible], Created)
DT   01-NOV-1999 (TrEMBLrel. [illegible], Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. [illegible], Last annotation update)
DE   HYPOTHETICAL [illegible] KDA PROTEIN
OS   Homo sapiens (Human)
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; [illegible]
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Rhodes S.;
RL   Submitted (MAY-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL049[illegible]; CAB4[illegible]
DR   InterPro; IPR00[illegible];
DR   InterPro; IPR00[illegible];
DR   Pfam; PF0062[illegible]
DR   Pfam; PF0185[illegible] START[illegible]
DR   SMART; SM0023[illegible] START[illegible]
KW   Hypothetical protein
SQ   SEQUENCE [illegible] AA; [illegible] MW; [illegible]

  Query Match [illegible]
  Best Local Similarity [illegible] Pred. No. [illegible]
  Matches [illegible] Conservative [illegible] Mismatches [illegible] Gaps [illegible]

Qy    59 [illegible]
Db   777 [illegible]
Qy   114 PLVG[illegible]KVE[illegible]EVV[illegible]EMER[illegible]VEEMEAM[illegible]WNENV[illegible]
Db   826 VGDGNP[illegible]KEW[illegible]ASVE[illegible]VVENEVEREGE[illegible]
Qy   16[illegible] E[illegible]LAAEAA[illegible]V[illegible]
Db   881 YV[illegible]MAP[illegible]EVVERIWE[illegible]
Qy   22[illegible] P[illegible]MVLHE[illegible]KTE[illegible]
Db   933 [illegible]

RESULT 14
O14868
ID   O14868     PRELIMINARY; [illegible] AA.
AC   O14868;
DT   01-JAN-1998 (TrEMBLrel. [illegible], Created)
DT   01-JAN-1998 (TrEMBLrel. [illegible], Last sequence update)
DT   [illegible] MAR-[illegible] (TrEMBLrel. [illegible], Last annotation update)
DE   HP PROTEIN
```

```
OS   [illegible]
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; [illegible]
OC   Mammalia; Eutheria; [illegible]
OX   NCBI_TaxID=[illegible]
RN   [illegible]
RP   SEQUENCE FROM N.A.
RA   [illegible]
RL   Submitted [illegible] EMBL/GenBank/DDBJ databases.
DR   EMBL; A[illegible]
[illegible]
SQ   SEQUENCE [illegible] MW; [illegible]

  Query Match [illegible]
  Best Local Similarity [illegible] Pred. No. [illegible]
  Matches [illegible] Mismatches [illegible] Indels [illegible] Gaps [illegible]

Qy   [illegible]AVMA[illegible]NEEA[illegible]
Db   [illegible]
Qy   [illegible]
Db   [illegible]
Qy   [illegible]WKEE[illegible]KVMSKVVE[illegible]EVVVE[illegible]
Db   [illegible]KEVKEKE[illegible]SYKKV[illegible]AVE[illegible]
Qy   [illegible]
Db   [illegible]
Qy   [illegible]
Db   [illegible]
Qy   [illegible]
Db   [illegible]

RESULT [illegible]
[illegible]
ID   [illegible] AA.
AC   [illegible]
DT   [illegible] MAR-[illegible] (TrEMBLrel. [illegible], Created)
DT   [illegible] MAR-[illegible] (TrEMBLrel. [illegible], Last sequence update)
DT   [illegible] MAR-[illegible] (TrEMBLrel. [illegible], Last annotation update)
DE   [illegible]
GN   [illegible]
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; [illegible] Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [illegible]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=[illegible]
RA   [illegible]
RA   [illegible]
RT   "[illegible] phosphorylation
RT   [illegible]
RT   [illegible]"
RL   [illegible]
```

```
DR   EMBL; AF232934; AAC42050.1; -.
SQ   SEQUENCE   598 AA;  68283 MW;  8474A[illegible]5A1A73 [illegible]

  Query Match             7.4%;  Score 108.5;  DB 11;  Length 598;
  Best Local Similarity   19.0%;  Pred. No. 2.2[illegible]
  Matches   67;  Conservative   61;  Mismatches  129;  Indels   [illegible];  Gaps   [illegible];

Qy          2 LLATFKLC---      ACSSYRHMKNMKCLR  QLAVMAIRQELNER ALGGPT-   -PS 46
              :|||    |                 || |  ::    |     ::     : :|| ::     ||       ::
Db        255 ILATLSPCIELMVEKEESWQFFHDFLVLKGRRVTLTYKNVMSELKKKEPTGTCYLECEN 314

Qy         47 TWINQ-----------------VRRRSSLLGSRLE- ---  -ETLYSLQELAYLQQGEEA 82
              : ||:                 :   :|       ||           :|   |       : :  :   :
Db        315 SLINPEEEFDAVEAALDRQDKIEEQSQSEKVR.SWPTSLPSCPTPTCVTHETVQKVLEM 374

Qy         83 MQKALGILSNQFCWKKESQQDNGLKVMSEVVEPVQ -EVEKLEV- - VVEQIMEELY   134
              :|    :                  | || |       ::| :  |  :| ||     :|     : :   :
Db        375 VQNHMNY----------SLQDVGGDANWQLVVE GEMKVYRREVEENGIVLDPLKAIHAV 424

Qy        135 -----EELVE    EMEAMGEWNPNVEEIEVL[illegible]NLVGPKD 183
                     |:         ::     :|      ::     |:: :   :        ||:       :        ||
Db        425 KGVTGHEVCNYFWNVDVRNDWETTIENFHVVE LADNATIVYQTHKRVWPAS - --QRD 479

Qy        184 FVSVRCAKP--  ----RCSTCVLACMATLPGNMPI QKGVIRAEHGPT-MVLHPLAGSES- 235
                :  :    ::                 | ::      :  |     :             :||:             :      |
Db        480 VLYLSAIRKIPALTENDLETWIVCNFSVDHD[illegible] 538

Qy        236 ---------KTELIWLLSIDLKGWLPKSIINQVL LTQVDFANHLRKKLESH 278
                       |:|:: :::    || | |::                  |          ||     :
Db        539 DQFISRDNILCKITYVANVNPGGWAPASVLRAVA--KEYPKFL---KKFISY 596


Search completed: August 22, 2001, 14:[illegible]7:[illegible]
Job time: 205 sec
```

```
                          GenCore version 4.5
               Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on:          August 22, 2001, 14:29:12 ; Search time 101.[illegible] Seconds
                                             (without alignments)
                                             [illegible].457 Million cell updates/sec

Title:           US-08-659-254-18
Perfect score:   1470
Sequence:        1 MLLATFKLCAGSSYRHMRNM ........[illegible]  285

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters:       41[illegible]

[illegible]mum DB seq length: 0
[illegible]mum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       A_Geneseq_0601:*
                1:   /SIDS8/gcgdata/geneseq/[illegible]p/AA[illegible]:*
                2:   /SIDS8/gcgdata/geneseq/[illegible]p/AA[illegible]:*
                3:   /SIDS8/gcgdata/geneseq/[illegible]p/AA[illegible]:*
                4:   /SIDS8/gcgdata/geneseq/[illegible]p/AA[illegible]:*
                5:   /SIDS8/gcgdata/geneseq/geneseqp/AA1984.DAT:*
                6:   /SIDS8/gcgdata/geneseq/geneseqp/AA1985.DAT:*
                7:   /SIDS8/gcgdata/geneseq/geneseqp/AA1986.DAT:*
                8:   /SIDS8/gcgdata/geneseq/geneseqp/AA1987.DAT:*
                9:   /SIDS8/gcgdata/geneseq/geneseqp/AA[illegible]:*
               10:   /SIDS8/gcgdata/[illegible]p/AA1989.DAT:*
               11:   /SIDS8/gcgdata/[illegible]p/AA[illegible]:*
               12:   /SIDS8/gcgdata/[illegible]p/AA[illegible]:*
               13:   /SIDS8/gcgdata/[illegible]p/AA1992.DAT:*
               14:   /SIDS8/gcgdata/[illegible]p/AA[illegible]:*
               15:   /SIDS8/gcgdata/[illegible]p/AA[illegible]:*
               16:   /SIDS8/gcgdata/[illegible]p/AA[illegible]:*
               17:   /SIDS8/gcgdata/[illegible]p/AA[illegible]:*
               18:   /SIDS8/gcgdata/[illegible]p/AA[illegible]:*
               19:   /SIDS8/gcgdata/[illegible]p/AA[illegible]:*
               20:   /SIDS8/gcgdata/[illegible]p/AA[illegible]:*
               21:   /SIDS8/gcgdata/[illegible]p/AA[illegible]:*
               22:   /SIDS8/gcgdata/[illegible]p/AA[illegible]:*

    Pred. No. is the number of results predicted by chance to have a
    score greater than or equal to the score of the result being printed,
    and is derived by analysis of the total score distribution.

                              SUMMARIES
                %
Result          Query
  No.   Score   Match Length DB   ID                        Description
-----------------------------------------------------------------------------
   1     1470   100.0    285  22  AAY728[illegible]         Human steroidogenic
   2     1464    99.6    285  17  AAW[illegible]0814        Human steroidogene
   3     1460    99.3    285  17  AAW[illegible]            Human steroidogene
   4     1460    99.3    285  17  AAW[illegible]            Human steroidogene
   5     1457    99.1    285  17  AAW[illegible]            Human steroidogene
   6     1457    99.1    285  17  AAW[illegible]            Human steroidogene
   7     1457    99.1    285  17  AAW[illegible]            Human steroidogene
   8   1448.5    98.5    284  17  AAW[illegible]            Human steroidogene
   9   1291.5    87.9    284  20  AAW[illegible]            Mouse steroidogenic
  10   1291.5    87.9    284  22  AAY[illegible]            Mouse steroidogenic
  11      372    25.3    445  18  AAW[illegible]            Human MLN [illegible]
```

```
  12      303    20.6    180  21  AAB[illegible]            Human cancer assoc
  13      274    18.6     69  21  AAG017[illegible]         Human secreted pro
  14    180.5    12.3    2[illegible]  21  AAB[illegible]   Human [illegible]
  15      179    12.2    169  21  AAB4[illegible]           Human cancer assoc
  16      169    11.5    221  22  AAB2[illegible]           Human steroidogeni
  17    108.5     7.4    598  21  AAY[illegible]            Mouse Goodpasture
  18      108     7.3   1091  20  AAY3[illegible]           Protein encoded by
  19      105     7.1    242  21  AAG1[illegible]           Arabidopsis thalia
  20      105     7.1    253  21  AAG1[illegible]           Arabidopsis thalia
  21      104     7.1    232  21  AAG[illegible]            Arabidopsis thalia
  22      104     7.1    242  21  AAG[illegible]            Arabidopsis thalia
  23      104     7.1    9[illegible]  [illegible]  AAB[illegible]  Transposase from a
  24     99.5     6.8    624  21  AAY[illegible]            Mouse Goodpasture
  25     98.5     6.7    291  21  AAB28[illegible]          Protein expressed
  26     98.5     6.7    291  21  AAB2[illegible]           Protein expressed
  27     98.5     6.7    291  22  AAB[illegible]            Mouse GN protein s
  28     98.5     6.7    291  22  AAB[illegible]            Amino acid sequence
  29     98.5     6.7    347  21  AAB4[illegible]           Human cancer assoc
  30     94.5     6.4    596  21  AAY9[illegible]           Goodpasture antige
  31     94.5     6.4    616  21  AAY[illegible]            Goodpasture antige
  32     94.5     6.4    617  21  AAY[illegible]            Goodpasture antige
  33     94.5     6.4    624  21  AAB4[illegible]           Human [illegible]EX [illegible]
  34     94.5     6.4    624  21  AAY[illegible]            Human Goodpasture
  35     94.5     6.4    637  21  AAY[illegible]            Goodpasture antige
  36     94.5     6.4   3433  18  AAW22[illegible]          Utrophin, [illegible]
  37     93.5     6.4    389  21  AAB41[illegible]          Human [illegible]X [illegible]
  38     93.5     6.4    598  21  AAY[illegible]            Human Goodpasture
  39     93.5     6.4    598  21  AAY[illegible]            Cattle Goodpasture
  40     93.5     6.4    624  21  AAY[illegible]            Cattle Goodpasture
  41     90.5     6.2    494  20  AAY2814[illegible]        Bacillus lichenifo
  42       [illegible]    [illegible]  [illegible]  20  AAW71[illegible]   Mouse steroidogeni
  43       [illegible]    [illegible]  [illegible]  22  AAY728[illegible]  Mouse tryptic pept
  44     [illegible]     [illegible]  [illegible]  [illegible]  AAW[illegible]  [illegible]
  45     [illegible]     [illegible]  [illegible]  21  AA[illegible]  Arabidopsis thalia

                              ALIGNMENTS

RESULT    1
AAY72838
ID   AAY72838 standard; protein; 285 AA.
XX
AC   AAY72838;
XX
DT   31-MAY-2001  (first entry)
XX
DE   Human steroidogenic acute regulatory (StAR) protein.
XX
KW   Human; [illegible];
KW   [illegible];
KW   lipoid congenital adrenal hyperplasia; sexual maturation;
KW   infertility; androgen responsive [illegible]; precocious puberty;
KW   hypogonadotropic hypogonadism; McCune-Albright syndrome;
KW   adrenal hypoplasia congenita; cytostatic.
XX
OS   Homo sapiens.
XX
FH   Key                 Location/Qualifiers
FT   Misc difference     156
FT                       /note= "Encoded by [illegible] that [illegible] corresponds
FT                       to the translation [illegible] for the sequence shown
FT                       in (AAD02909)"
FT   Modified site       59..60
FT                       /note= "PKA phosphorylation site"
FT   Modified-site       68..69
FT                       /note= "CK phosphorylation site"
FT   Modified-site       [illegible]5
FT                       /note= "PKA phosphorylation site"
FT   Modified-site       23[illegible]..234
FT                       /note= "P34cDC2 kinase ([illegible]) phosphorylation site"
FT   Modified site       238
FT                       /note= "PKC phosphorylation site"
FT   Modified-site   [illegible]
FT                   /note= "PKA p[illegible]"
XX
PN   US6194555-B1.
XX
PD   27-FEB-2001.
XX
PF   07-JUN-1996; [illegible]
XX
PR   04-NOV-1995; [illegible]
XX
PA   (BYTE ) UNIV TEXAS [illegible]
XX
PI   Stocco DM, [illegible]
XX
DR   WPI; 2001-[illegible]
DR   N-PSDB; AAD[illegible] AA[illegible]
XX
PT   Novel nucleic acid encoding [illegible]
PT   useful for treating steroid hormone dependent [illegible]
XX
PS   Example 14; [illegible] English.
XX
CC   The invention relates to steroidogenic acute regulatory protein
CC   (StAR) and nucleic acid molecules encoding the [illegible]
CC   are used for regulating steroidogenesis. More [illegible]
CC   proteins are involved in regulating [illegible]
CC   the mitochondria of [illegible]
CC   invention are useful in gene therapy for the treatment [illegible]
CC   hormone dependent disorders [illegible]
CC   adrenal hyperplasia (CAH), infertility, [illegible]
CC   androgen responsive tumours, precocious puberty, McCune-Albright
CC   syndrome, adrenal hyp[illegible]
CC   hypogonadism. They are also useful [illegible]
CC   of the steroidogenic acute regulatory protein
CC   The present sequence [illegible]
CC   protein (StAR).
XX
SQ   Sequence 284 AA;

  Query Match   [illegible]
  Best Local Similarity [illegible]
  Matches  [illegible]

Qy          1 MLLATFKL[illegible]YRHMRNMK[illegible]AVMA[illegible]
Db          1 mllatfkl[illegible]

Qy         61 SRLEESTLYSDQ[illegible]AYL[illegible]EEAMQKA[illegible]
Db         61 [illegible]

Qy        121 RFEVVVD[illegible]EM[illegible]AM[illegible]WNPNVK[illegible]
Db        121 [illegible]

Qy        181 PRDFVSVR[illegible]AK[illegible]ST[illegible]AGMA[illegible]NM[illegible]
Db        181 pr[illegible]vsvr[illegible]st[illegible]

Qy        241 WLL[illegible]DLKGWL[illegible]VL[illegible]QVI[illegible]ANHLRK[illegible]
Db        241 wll[illegible]dlkgwl[illegible]


RESULT   2
AAW[illegible]
ID   AAW[illegible] standard; [illegible] 285 AA.
XX
AC   AAW[illegible];
XX
DT   21-MAY-1997  (first entry)
```

```
XX
DE   [illegible]
XX
KW   [illegible] StAR [illegible]
KW   [illegible]
KW   [illegible] treatment [illegible]
KW   [illegible]
XX
OS   [illegible]
XX
FH   Key             Location/Qualifiers
FT   [illegible]
FT                   /note= "potential site for protein kinase A
FT                   mediated phosphorylation"
FT   [illegible]
FT                   /note= "potential site for protein kinase A
FT                   mediated phosphorylation"
FT   [illegible]
FT                   /note= "potential site for protein kinase A
FT                   mediated phosphorylation"
FT   [illegible]
FT                   /note= "potential site for protein kinase A
FT                   mediated phosphorylation"
FT   [illegible]
FT                   /note= "potential site for protein kinase A
FT                   mediated phosphorylation"
FT   [illegible]
FT                   /note= "potential site for protein kinase A
FT                   mediated phosphorylation"
FT   [illegible]
FT                   /note= "potential site for protein kinase A
FT                   mediated phosphorylation"
FT   [illegible]
FT                   /note= "potential site for protein kinase A
FT                   mediated phosphorylation"
XX
PN   W[illegible]
XX
PD   [illegible]
XX
PF   [illegible] MAR [illegible]
XX
PR   [illegible] MAR [illegible]
XX
PA   (REGC ) UNIV CALIFORNIA.
PA   (UYPE ) UNIV PENNSYLVANIA.
XX
PI   [illegible]
XX
DR   W[illegible]
DR   N-PSDB; AA[illegible]
XX
PT   [illegible]
PT   [illegible]
PT   [illegible]
XX
PS   [illegible]
XX
CC   The present sequence [illegible]
CC   regulatory protein (hStAR). The hStAR gene can be analysed [illegible]
CC   [illegible] (e.g. prenatally) genetic defects associated
CC   with [illegible] adrenal hyperplasia (CAH) [illegible]
CC   [illegible] StAR can be treated [illegible]
CC   [illegible] therapy [illegible] present [illegible]
CC   [illegible]
CC   [illegible] DNA [illegible] StAR
CC   [illegible]
CC   [illegible]
CC   [illegible] transfected with [illegible] StAR [illegible] increased [illegible]
CC   [illegible]
CC   [illegible]
XX
SQ   Sequence [illegible] AA;
```

```
  Query Match             99.6%;  Score 146[illegible];  DB 17;  Length [illegible];
  Best Local Similarity   99.6%;  Pred. No. [illegible]-142;
  Matches  284;  Conservative    0;  Mismatches    1;  Indels  [illegible];  Gaps  [illegible];

Qy          1 MLLAIFKLCAGSSYRHMRNMKGLRQQAVMA[illegible]
              ||||||||||||||||||||||||||||||[illegible]
Db          1 mllatfklcaqssyrhmrnmkqlrqqavmais[illegible]

Qy         61 SRLEETLYSDQELAYLQQGEEAMQKA[illegible]
              ||||||||||||||||||||||||||[illegible]
Db         61 srleetlysdqelaylqqgeeamqkal[illegible]

Qy        121 R[illegible]
              ||||||||||||||||||||||||||[illegible]
Db        121 r[illegible]qpmer[illegible]

Qy        181 PRDFVSVRCAKRR[illegible]VLAGMA[illegible]SPSKTKLI 24[illegible]
              ||||||||||||[illegible]
Db        181 prdfvsvrcak[illegible]

          241 WLLSIDLKGWLPKSIINQVLSQTQVDFANHLRK[illegible]LESHPASEARC 285
              ||||||||||||||||||||||||||||||||[illegible]
D         241 wllsidlk[illegible]

RESULT    3
AAW13236
ID   AAW13236 standard; Protein; 285 AA.
XX
AC   AAW13236;
XX
DT   21-MAY-1997  (first entry)
XX
DE   Human steroidogenesis acute regulator, protein mutant Ala218Val.
XX
KW   Human; steroidogenesis; acute regulat[illegible] protein; hStAR; analysis;
KW   mutation; detection, prenatal, genetic defect; congenital; protein;
KW   lipoid adrenal hyperplasia; treatment; prevention; gene;
KW   replacement therapy; hypercholesterolaemia; mutant.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 218
FT                   /note= "wild type Al[illegible] substituted with Val"
XX
PN   WO9629338-A1.
XX
     26-SEP-1996.

     22-MAR-1996;    96WO-US03896
XX
PR   23-MAR-1995;    95US-0410540.
XX
PA   (REGC ) UNIV CALIFORNIA.
PA   (UYPE-) UNIV PENNSYLVANIA.
XX
PI   Lin D,  Miller WL,  Strauss JF;
XX
DR   WPI; 1996-443130/44.
DR   N PSDB; AAT62021.
XX
PT   Isolated human steroidogenesis acute regulatory protein gene - used
PT   for detection of mutation(s) of this gene that cause congenital
PT   lipoid adrenal hyperplasia
XX
PS   Claim 13; Page -; 89pp; English.
XX
CC   The present sequence is the human steroidogenesis acute regulatory
CC   protein (hStAR), congenital lipoid adrenal hyperplasia (CAH)
CC   associated mutant, Ala218Val. Analysing the hStAR gene for such a
```

```
CC   mutation is useful in the detection (e.g. prenatally) of CAH
CC   associated genetic defects, or its [illegible] to children. CAH
CC   can be treated by protein or gene replacement therapy, which can
CC   also be used to prevent or treat hypercholesterolaemia.
XX
SQ   Sequence   285 AA;

  Query Match             99.[illegible]%;  Score [illegible];  DB [illegible];  Length [illegible];
  Best Local Similarity   99.3%;  Pred. No. [illegible]-142;
  Matches  283;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 MLLAIFKL[illegible]
              ||||||||[illegible]
Db          1 mllatfkl[illegible]rqqavm[illegible]wln[illegible]vrrssll[illegible] 6[illegible]

Qy         61 SRLEETL[illegible]ELAYL[illegible]EAMQKA[illegible]
              ||||||||||||[illegible]
Db         61 srleetlysdqelaylqqgeeamqkal[illegible]

Qy        121 R[illegible]LVVV[illegible] 180
              |||||||||[illegible]
Db        121 rlevvvdqpmerlyee[illegible] 180

Qy        181 PRDFVSVR[illegible]VLAGMA[illegible]NM[illegible] 240
              ||||||||||||||||||||||[illegible]
Db        181 prdfvsvrcakrrgstcvlagmdtdlqnm[illegible]laaspsktkli 240

Qy        241 WLLSIDLKGWLPKSIINQVLSQTQVDFANHLRK[illegible]ARC 285
              |||||||||||||||||||||||||||||[illegible]
Db        241 wllsidlkgwlpksiinqvlsqtqvdfanhlrk[illegible]arc 285

RESULT    4
AAW13235
ID   AAW13235 standard; Protein; 285 AA.
XX
AC   AAW13235;
XX
DT   21-MAY-1997  (first entry)
XX
DE   Human steroidogenesis acute regulatory protein mutant Glu169Lys.
XX
KW   Human; steroidogenesis; acute regulatory protein; hStAR; analysis;
KW   mutation; detection; prenatal; genetic defect; congenital; protein;
KW   lipoid adrenal hyperplasia; treatment; prevention; gene;
KW   replacement therapy, hypercholesterolaemia; mutant.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 169
FT                   /note= "wild type [illegible] substituted with Lys"
XX
PN   WO9629338-A1.
XX
PD   26-SEP-1996.
XX
PF   22-MAR-1996;    96WO-US03896.
XX
PR   23-MAR-1995;    95US-0410540.
XX
PA   (REGC ) UNIV CALIFORNIA.
PA   (UYPE ) UNIV PENNSYLVANIA.
XX
PI   Lin D,  Miller WL,  Strauss JF;
XX
DR   WPI; 1996-443130/44.
DR   N-PSDB; AAT62020.
XX
PT   Isolated human steroidogenesis acute regulatory protein gene - used
PT   for detection of mutation(s) of this gene that cause congenital
PT   lipoid adrenal hyperplasia
```

```
XX
PS   Claim 13; Page [illegible]; [illegible] English.
XX
CC   The present sequence [illegible]
CC   protein (hStAR) [illegible] congenital lipoid adrenal [illegible]
CC   associated mutant, [illegible] StAR [illegible]
CC   mutation is useful [illegible] the detection [illegible]
CC   associated genetic defects [illegible]
CC   can be treated by [illegible] replacement [illegible]
CC   also be used [illegible]
XX
SQ   Sequence [illegible] AA;

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches 283; Conservative [illegible] Mismatches [illegible]

Qy          1 M[illegible]LATFK[illegible]A[illegible]SYE[illegible]MENMK[illegible]LAVMA[illegible]NFRA[illegible]
Db          1 ml[illegible]atfk[illegible]syr[illegible]menmkql[illegible]
Qy         61 [illegible]RLEEEILYS[illegible]AY[illegible]EEAMQKA[illegible]NQ[illegible]WKEE[illegible]
Db         61 [illegible]eetly[illegible]
Qy        121 [illegible]LEVVVD[illegible]M[illegible]LYE[illegible]AM[illegible]NVK[illegible]
Db        121 [illegible]
Qy        181 P[illegible]FVSVR[illegible]GS[illegible]VLA[illegible]MA[illegible]NMF[illegible]K[illegible]VLFAH[illegible]
Db        181 p[illegible]
Qy        241 W[illegible]K[illegible]W[illegible]KSF[illegible]
Db        241 w[illegible]

RESULT [illegible]
AAW1[illegible]
ID   AAW1[illegible] standard; Protein; [illegible] AA.
XX
AC   AAW1[illegible]
XX
DT   21 MAY 1997 (first entry)
XX
DE   Human steroidogenic [illegible]
XX
KW   Human; steroidogenesis; [illegible]
KW   mutation; detection; [illegible]
KW   lipoid adrenal hyperplasia; treatment; [illegible]
KW   replacement therapy; [illegible]
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference [illegible]
FT                   /note= "wild type [illegible]"
XX
PN   WO9629[illegible] A1.
XX
PD   26-SEP-1996.
XX
PF   22-MAR-1996; [illegible]
XX
PR   23-MAR-1995; [illegible]
XX
PA   (REGC ) UNIV CALIFORNIA.
PA   (TRUS ) UNIV PENNSYLVANIA.
XX
PI   Lin D, Miller WL, [illegible]
XX
```

```
[illegible]
XX
CC   [illegible]
XX
SQ   Sequence [illegible] AA;

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches [illegible] Mismatches [illegible] Indels [illegible] Gaps [illegible]

[illegible]

RESULT [illegible]
AAW1[illegible]
ID   AAW1[illegible] AA.
XX
AC   AAW1[illegible]
XX
DT   [illegible] MAY [illegible]
XX
DE   Human [illegible]
XX
KW   [illegible]
KW   [illegible]
KW   [illegible]
KW   [illegible]
XX
OS   [illegible]
XX
FH   [illegible]
FT   Misc-difference [illegible]
FT   [illegible]
XX
PN   W[illegible]
XX
PD   [illegible]
XX
PF   [illegible] MAR [illegible]
XX
PR   [illegible] MAR [illegible]
```

```
XX
PA   (REGC ) UNIV CALIFORNIA.
PA   (UYPE-) UNIV PENNSYLVANIA.
XX
PI   Lin D,  Miller WL,  Strauss JF;
XX
DR   WPI; 1996-443130/44
DR   N PSDB; AAT62018.
XX
PT   Isolated human steroidogenesis acute regulatory protein gene - used
PT   for detection of mutation(s) of this gene that cause congenital
PT   lipoid adrenal hyperplasia
XX
PS   Claim 13; Page -; 89pp; English.
XX
CC   The present sequence is the human steroidogenesis acute regulatory
CC   protein (hStAR), congenital lipoid adrenal hyperplasia (CAH)
CC   associated mutant, Glu169Gly. Analysing the hStAR gene for such a
CC   mutation is useful in the detection (e.g. prenatally) of CAH
CC   associated genetic defects, or its transmission to children. CAH
CC   can be treated by protein or gene replacement therapy, which can
CC   also be used to prevent or treat hypercholesterolaemia.
[illegible]
[illegible]  Sequence   285 AA;

  Query Match             99.1%;  Score 14[illegible];  DB 17;  Length 285;
  Best Local Similarity   99.3%;  Pred. No. [illegible] 7e-141;
  Matches  283,  Conservative    0,  Mismatches    2,  Indels  [illegible],  Gaps  [illegible],

Qy        1 MLLATFKLCAGSSYRHMRNMKGLRQQAVMAISQELNRRALGGPTPSTWINQVRRRSSLL[illegible] 60
            [illegible]
Db        1 mllatfklcagssyrhmrnmkglrqqavma[illegible] 60

Qy       61 SRLEETLYSDQELAYLQQGEEAMQKALGILSN[illegible] 120
            [illegible]
Db       61 srleetlysdqelaylqqgeeamqkal[illegible]

Qy      121 RLEVVVDQPMERLYEELVERMEAMGEWNPNVK[illegible]
            [illegible]
Db      121 [illegible]

Qy      181 PRDFVSVRCAKRRGSTCVLAGMATDFGNMPE[illegible] 240
            [illegible]
Db      181 prdfvsvrcakrrgst[illegible] 240

Qy      241 WLLSIDLKGWLPKSIINQVLSQTQVDFANHLRKRLESHPASEARC 285
            [illegible]
Db      241 wllsidlkgw[illegible] 2[illegible]


[illegible]LT    7
[illegible]3234
ID   AAW13234 standard; Protein; 285 AA.
XX
AC   AAW13234;
XX
DT   21-MAY-1997  (first entry)
XX
DE   Human steroidogenesis acute regulatory protein mutant Arg182Leu.
XX
KW   Human; steroidogenesis; acute regulatory protein; hStAR; analysis;
KW   mutation; detection; prenatal; genetic defect; congenital; protein;
KW   lipoid adrenal hyperplasia; treatment; prevention; gene;
KW   replacement therapy; hypercholesterolaemia; mutant.
XX
OS   Homo sapiens.
XX
FH   Key                 Location/Qualifiers
FT   Misc difference 182
FT                       /note= "wild type Arg substituted with Leu"
XX
PN   WO9629338-A1.
```

```
XX
PD   26-SEP-1996.
XX
PF   22-MAR-1996;    96WO-US03896.
XX
PR   23-MAR-1995;    95US-0410540.
XX
PA   (REGC ) UNIV CALIFORNIA.
PA   (UYPE-) UNIV PENNSYLVANIA.
XX
PI   Lin D,  Miller WL,  Strauss JF;
XX
DR   WPI; 1996-443[illegible]/44.
DR   N-PSDB; AAT620[illegible].
XX
PT   Isolated human steroidogenesis acute regulatory protein gene - used
PT   for detection of mutation(s) of this gene that cause congenital
PT   lipoid adrenal hyperplasia
XX
PS   Claim 13; Page -; 89pp; English.
XX
CC   The present sequence is the human steroidogenesis acute regulatory
CC   protein (hStAR), congenital lipoid adrenal hyperplasia (CAH)
CC   associated mutant, Arg182Leu. Analysing the hStAR gene for such a
CC   mutation is useful in the detection (e.g. prenatally) of CAH
CC   associated genetic defects, or its transmission to children. CAH
CC   can be treated by protein or gene replacement therapy, which can
CC   also be used to prevent or treat hypercholesterolaemia.
XX
SQ   Sequence   285 AA;

  Query Match             99.1%;  Score 1[illegible];  DB 17;  Length 285;
  Best Local Similarity   99.3%;  Pred. No. [illegible];
  Matches  283,  Conservative    0,  Mismatches    2;  Indels    0;  Gaps    0;

Qy        1 MLLATF[illegible]NVRRRSSLL[illegible]
            [illegible]
Db        1 mllatfkl[illegible] 60

Qy       61 SRLEETLYSDQELAYLQQGEEAMQKAL[illegible]KVVPDVGKVF 120
            [illegible]
Db       61 srleetlysdqelaylqqgeeamqkal[illegible] 120

Qy      121 RLEVVV[illegible]LVG 180
            [illegible]
Db      121 rlevvvdqpme[illegible] 180

Qy      181 [illegible]VLAGMA[illegible]SKTKLT 240
            [illegible]
Db      181 pldfvsvrcakrrgstcvlagmatdfgnmpeqkgv[illegible] 240

Qy      241 WLLSIDLKGWL[illegible] 285
            [illegible]
Db      241 wllsidlkgwlpksi[illegible] 285


RESULT    8
AAW13238
ID   AAW13238 standard; Protein; 281 AA
XX
AC   AAW13238;
XX
DT   21-MAY-1997  (first entry)
XX
DE   Human steroidogenesis acute regulatory protein mutant [illegible]
XX
KW   Human; steroidogenesis; acute regulatory protein; hStAR; analysis;
KW   mutation; detection; prenatal; genetic defect; congenital; protein;
KW   lipoid adrenal hyperplasia; treatment; prevention; gene;
KW   replacement therapy; hypercholesterolaemia; mutant.
XX
OS   Homo sapiens.
```

```
XX
PN   W[illegible]08-A1
XX
PD   26-SEP-1996.
XX
PF   2[illegible]-MAR-1996; [illegible]
XX
PR   23-MAR-1995; [illegible]
XX
PA   (REGC ) UNIV CALIFORNIA.
PA   (UYPE ) UNIV PENNSYLVANIA.
XX
PI   Lin D, Miller WL, Strauss JF;
XX
DR   WPI; 1996-44[illegible]44.
XX
PT   Isolated human steroidogenesis acute regulatory [illegible]
PT   for detection of mutations [illegible]
PT   lipoid adrenal hyperplasia
XX
PS   Claim [illegible]; Page [illegible]; English.
XX
CC   The present sequence is the human [illegible]
CC   protein (hStAR) [illegible] lipoid adrenal [illegible]
CC   associated mutant [illegible]
CC   been deleted. [illegible]
CC   useful in the detection [illegible]
CC   genetic defects [illegible]
CC   treated by [illegible]
CC   used to prevent [illegible]
XX
SQ   Sequence [illegible] AA;

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches [illegible]

Qy          1 MLLATFKL[illegible]AVMA[illegible]NRRAL[illegible]
Db          1 mllatfkl[illegible]
Qy         61 SRLEETLY[illegible]AYL[illegible]
Db         61 [illegible]
Qy        121 REEVVVDQPMEAYEE[illegible]
Db        121 [illegible]vvdqpme[illegible]
Qy        181 PRRTVSVR[illegible]
Db        181 [illegible]
Qy        241 WLL[illegible]DLK[illegible]
Db        241 wll[illegible]dlk[illegible]


RESULT [illegible]
AAW[illegible]
ID   AAW[illegible] standard; [illegible] AA.
XX
AC   AAW[illegible];
XX
DT   [illegible]-MAY-[illegible] (first entry)
XX
DE   Mouse steroidogenic acute regulatory protein [illegible]
XX
KW   StAR; steroidogenic acute regulatory protein; [illegible]
KW   steroid hormone; [illegible]
KW   lipoid congenital adrenal hyperplasia; [illegible]
KW   androgen responsive; [illegible]
```

```
XX
[illegible]
XX
PN   [illegible]
XX
PD   [illegible]
XX
PF   [illegible]
XX
PR   [illegible]
XX
PA   [illegible]
XX
PI   [illegible]
XX
DR   [illegible]
DR   [illegible]
XX
PT   [illegible]
PT   [illegible]
PT   [illegible]
PT   [illegible]
XX
PS   [illegible]
XX
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
XX
SQ   Sequence [illegible] AA;

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches [illegible]

Qy   [illegible]
Db   [illegible]
Qy   [illegible]
Db   [illegible]
Qy   [illegible]
Db   [illegible]
Qy   [illegible]
Db   [illegible]
Qy   [illegible]
Db   [illegible]


RESULT [illegible]
AAW[illegible]
ID   [illegible] AA.
XX
AC   [illegible]
XX
DT   [illegible]-MAY-[illegible]
XX
DE   [illegible] (StAR [illegible]
XX
```

```
KW   Mouse; steroidogenic acute regulatory protein; StAR;
KW   gene therapy; steroid hormone dependent disorder; LCAH;
KW   lipoid congenital adrenal hyperplasia; sexual maturation;
KW   infertility; androgen responsive tumour; precocious puberty;
KW   hypogonadotropic hypogonadism; McCune Albright syndrome;
KW   adrenal hypoplasia congenita; cytostatic.
XX
OS   Mus sp.
XX
FH   Key                 Location/Qualifiers
FT   Peptide             1..47
FT                       /label= Signal_peptide
FT   Cleavage-site       39..40
FT                       /note= "Site for mitochondrial protease
FT                       matrix-processing protease and mitochondrial
FT                       intermediate processing peptide"
FT   Protein             48..284
FT                       /label= Mature_steroidogenic_acute_regulatory_protein
FT   Cleavage-site       49..50
FT                       /note= "Site for mitochondrial protease
FT                       matrix-processing protease and mitochondrial
FT                       intermediate processing peptide"
     Modified-site       59..60
                         /note= "PKA phosphorylation site"
     Modified-site       68..69
FT                       /note= "CK phosphorylation site"
FT   Region              85..96
FT                       /label= Tryptic_peptide_#[illegible]
FT   Region              88..98
FT                       /note= "This sequence is used to raise antibodies
FT                       against StAR protein"
FT   Region              193..211
FT                       /label= Tryptic_peptide_#[illegible]
FT   Modified-site       195
FT                       /note= "PKA phosphorylation site"
FT   Region              217..230
FT                       /label= Tryptic_peptide_#2[illegible]
FT   Modified-site       233..234
FT                       /note= "P34CDC2 kinase (CK) phosphorylation site"
FT   Modified-site       238
FT                       /note= "PKC phosphorylation site"
XX
PN   US6194555-B1.
XX
PD   27-FEB-2001.
XX
PF   07-JUN-1996;   96US-0659254.
XX
PR   04-NOV-1995;   95US-0538960.
XX
     (UYTE-) UNIV TEXAS TECH HEALTH SCI CEN
XX
     Stocco DM, Clark BJ;
XX
DR   WPI; 2001-217936/22.
DR   N-PSDB; AAD02901, AAD02902.
XX
PT   Novel nucleic acid encoding steroidogenic acute regulatory protein,
PT   useful for treating steroid hormone dependent disorders
XX
PS   Example 2; Fig 2; 41pp; English.
XX
CC   The invention relates to steroidogenic acute regulatory protein
CC   (StAR) and nucleic acid molecules encoding them. StAR proteins
CC   are used for regulating steroidogenesis. More particularly these
CC   proteins are involved in regulating cholesterol transport into
CC   the mitochondria of a cell. The nucleic acid molecules of the
CC   invention are useful in gene therapy or the treatment of steroid
CC   hormone-dependent disorders, particularly, lipoid congenital
CC   adrenal hyperplasia (LCAH), infertility, sexual maturation,
CC   androgen-responsive tumours, precocious puberty, McCune Albright
CC   syndrome, adrenal-hypoplasia congenita or hypogonadotropic
CC   hypogonadism. They are also useful for recombinant preparation
CC   of the steroidogenic acute regulatory protein.
CC   The present sequence is mouse steroidogenic acute regulatory
CC   protein (StAR).
XX
SQ   Sequence   284 AA

  Query Match              87.9%;  Score [illegible];  DB 12;  Length 284;
  Best Local Similarity    87.0%;  Pred. No. [illegible];
  Matches  248;  Conservative  19;  Mismatches  17;  Indels  1;  Gaps  1;

Qy          1 MLLATFKL[illegible]VRRSSL[illegible]
              [illegible]
Db          1 mllatfkle[illegible]
                
Qy         61 SRLEETLYS[illegible]AYL[illegible]EAM[illegible]KAL[illegible]VM[illegible]KVV[illegible]VGKVF 120
              [illegible]
Db         60 [illegible]lskrvpdvqkvf 11[illegible]

Qy        121 RLEVVV[illegible]EAMGEWNPNV[illegible]AAAAAA[illegible]
              [illegible]
Db        120 rlevvvdqpmdrlyeelvdrmeamgewnp[illegible]aaaaaa[illegible]

Qy        181 [illegible]VLAGMA[illegible]KIK[illegible] 240
              [illegible]
Db        180 prdfvsv[illegible]krrgst[illegible]vlaqmat[illegible]sktk[illegible] 239

Qy        241 W[illegible]EAF 285
              [illegible]
Db        240 wllsidlk[illegible]lpk[illegible]inqvlsqtqie[illegible]aseaq[illegible] 284


RESULT 11
AAW25768
ID   AAW25768 standard; Protein; 445 AA.
XX
AC   AAW25768;
XX
DT   05-DEC-1997  (first entry)
XX
DE   Human MLN 64.
XX
KW   Human; CART1; Lasp 1; MLN 64; MLN 5[illegible]; breast carcinoma cell;
KW   Cys-rich region; RING finger domain; CART motif; TRAF domain; LIM domain;
KW   TNF related cytokine signal transduction; breast cancer; SH3;
KW   Src homology region 3.
XX
OS   Homo sapiens.
XX
FH   Key                 Location/Qualifiers
FT   Domain              1..72
FT                       /label= Potential_transmembrane_domain
FT   Domain              94..168
FT                       /label= Potential_transmembrane_domain
XX
PN   WO9706256-A2.
XX
PD   20-FEB-1997.
XX
PF   31-JUL-1996;   96WO-US12560.
XX
PR   09-AUG-1995;   95US-0002183.
XX
PA   (BRIM ) BRISTOL MYERS SQUIBB CO.
PA   (CNRS ) CENT NAT RECH SCI.
PA   (INRM ) INST NAT SANTE & RECH MED CALE
PA   (UYPA-) UNIV PASTEUR LOUIS.
XX
PI   Basset P, Rio M, Tomasetto C;
XX
DR   WPI; 1997-154243/14.
DR   N-PSDB; AAT86[illegible].
XX
```

```
PT   Genes amplified and over expressed [illegible]
PT   chromosome [illegible]
PT   prognosis and [illegible] leukaemia markers
XX
PS   Claim 11; [illegible]
XX
CC   This sequence [illegible] MEN 64 encoded by DNA [illegible]
CC   The MEN 64 gene has been mapped [illegible]
CC   of chromosome [illegible] with the maximum [illegible]
CC   contains two [illegible] MEN 64
CC   gene contains [illegible] exons. There are [illegible] MEN 64 [illegible]
CC   from nucleotide [illegible]
CC   modifications [illegible]
CC   the MEN 64 variants result from [illegible]
CC   variants lead to amino acid variant [illegible]
CC   C truncated MEN 64 proteins. A [illegible]
CC   containing [illegible]
CC   residues. The full length MEN 64 [illegible]
CC   transmembrane domain and [illegible]
CC   leucine rich repeat [illegible] MEN 64 [illegible]
CC   useful as prognostic markers [illegible]
XX
SQ   Sequence 314 AA;

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches 78; Conservative 49; Mismatches [illegible]

Qy    18 YSSQELAY[illegible]KAL[illegible]SWKKL[illegible]FVM[illegible]
              |              |
Db   231 [illegible]

Qy   128 QPMERLYEELV[illegible]RMF[illegible]WNPNVF[illegible]
Db   291 [illegible]

Qy   198 F[illegible]AKRR[illegible]V[illegible]A[illegible]MA[illegible]NMPE[illegible]
               ||         |
Db   351 [illegible]

Qy   248 KGWLPKS[illegible]N[illegible]VFEANH[illegible]
Db   411 k[illegible]prvl[illegible]

RESULT 12
AAB43728
ID   AAB43728 standard; Protein; [illegible] AA.
XX
AC   AAB43728;
XX
DT   [illegible] FEB [illegible]
XX
DE   Human cancer [illegible]
XX
KW   Human; cancer [illegible] gene [illegible]
KW   diagnosis; cytostatic; proliferation [illegible]
KW   antidiabetic; antiasthmatic; antirheumatic [illegible]
KW   antiinflammatory; antithyroid; antibacterial [illegible]
KW   dermatological; [illegible]
KW   gastrointestinal [illegible]
KW   immune disorder [illegible]
KW   allergic reaction [illegible]
KW   haemostatic; [illegible]
KW   neurological disease; [illegible]
XX
OS   Homo sapiens.
XX
PN   WO200105[illegible]-A1
XX
PD   [illegible]-SEP-200[illegible]
XX
```

```
PF   [illegible]-MAR-[illegible]
XX
PR   [illegible]-MAR-[illegible]
XX
PA   (HUMA-) HUMAN GEN[illegible]
XX
PI   [illegible]
XX
DR   WPI; [illegible]
DR   N-PSDB; [illegible]
XX
PT   N[illegible]
PT   [illegible]
XX
PS   [illegible]
XX
CC   AA[illegible] [illegible] cancer associated proteins [illegible]
CC   [illegible] AAB[illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible] antiasthmatic, antirheumatic, antiarthritic,
CC   [illegible] antibacterial [illegible]
CC   [illegible] embodiments [illegible]
CC   [illegible] the
CC   [illegible] present [illegible]
CC   [illegible]
CC   [illegible]
CC   the present invention [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible] disease and [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible]
CC   [illegible] AAB[illegible]
CC   AA[illegible] sequence [illegible]
CC   the present [illegible]
XX
SQ   [illegible]

  Query Match [illegible]
  Best Local Similarity [illegible]; Pred. No. [illegible]
  Matches [illegible]; Mismatches [illegible]; Indels [illegible]; Gaps [illegible]

Qy   [illegible]FVM[illegible]KVV[illegible]MERLYEE[illegible]WNPNVF[illegible]K[illegible] 104
Db   [illegible]

Qy   [illegible]AAFAAN[illegible]V[illegible]AKRR[illegible]NMP[illegible]KAL[illegible] 154
Db   [illegible]

Qy   [illegible]
Db   [illegible]

RESULT [illegible]
AA[illegible]
ID   AA[illegible] AA.
XX
AC   AA[illegible]
XX
DT   [illegible]
XX
DE   [illegible]
XX
KW   [illegible]NA[illegible]
KW   [illegible]
XX
OS   Homo sapiens.
XX
```

```
PN   EP1033401-A2.
XX
PD   06-SEP-2000.
XX
PF   21-FEB-2000; 2000EP-0200610.
XX
PR   26-FEB-1999;   99US-0122487.
XX
PA   (GEST ) GENSET.
XX
PI   Dumas Milne Edwards J,  Duclert A,  Giordano J;
XX
DR   WPI; 2000-500381/45.
DR   N-PSDB; AAC01736.
XX
PT   New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT   obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT   diagnostic, forensic, gene therapy and chromosome mapping procedures
XX
PS   Claim 13; SEQ ID 5811; 71pp + CD-ROM; English.
XX
CC   The present sequence is a polypeptide encoded by one of a large number
     of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs
     were prepared from total human RNAs or polyA+ RNAs derived from 29
     different tissues. EST sequences usually correspond mainly to the 3'
CC   untranslated region (UTR) of the mRNA because they are often obtained
CC   from oligo-dT primed cDNA libraries. Such ESTs are not well suited for
CC   isolating cDNA sequences derived from the 5' ends of mRNAs and even in
CC   those cases where longer cDNA sequences have been obtained, the full 5'
CC   UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'
CC   ends and can therefore be used to obtain full length cDNAs and genomic
CC   DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and
CC   chromosome mapping procedures. They are used to obtain upstream
CC   regulatory sequences and to design expression and secretion vectors.
XX
SQ   Sequence   69 AA;


  Query Match             18.6%;  Score 271;  DB 21;  Length 69;
  Best Local Similarity   91.9%;  Pred. No. [illegible].1e-21;
  Matches   57;  Conservative    1;  Mismatches    2;  Indels    2;  Gaps    1;

Qy        1 MLLATPKLCAGSSYRHMRNMKCLRQQAVMA[illegible]NRRALGGPTPS[illegible]NQVI[illegible]RS  SL 58
            ||||||||||||||||||||||||||||||||     |   ||| |||||| |  |  |         :
Db        1 mllatrklcaqssyrhmrnmkqlrqqavmals[illegible]rrdlqgptpr[illegible]xlysv 60

Qy       59 LG 60
            ||
Db       61 lq 62


[illegible]LT   14
[illegible]3294
     AAB53294 standard; Protein; 239 AA.
XX
AC   AAB53294;
XX
DT   09-MAR-2001  (first entry)
XX
DE   Human colon cancer antigen protein sequence SEQ ID No:834.
XX
KW   Human; colon cancer; colon cancer antigen; diagnosis; detection;
KW   identification; cytostatic; cardioactive; neuroprotective; vulnerary;
KW   immunomodulatory, muscular, gynaecological, gastrointestinal;
KW   nephrotropic; antiinfective; antibacterial; gene therapy; wound;
KW   neural disorder; immune system disorder; muscular disorder;
KW   reproductive disorder; gastrointestinal disorder; renal disorder;
KW   infectious disease; cardiovascular disorder.
XX
OS   Homo sapiens.
XX
PN   WO200055351-A1.
XX
PD   21-SEP-2000.
XX
PF   08-MAR-2000; 2000WO-US05883.
XX
PR   12-MAR-1999;   99US-0124270.
XX
PA   (HUMA ) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Ruben SM;
XX
DR   WPI; 2000-587541/55.
DR   N-PSDB; AAC98031.
XX
PT   Colon cancer associated gene sequences - referred to as colon cancer
PT   antigens, useful for the treatment, prevention, and diagnosis of colon
PT   disorders such as colon cancer
XX
PS   Claim 11; Page 1385; 2104pp; English.
XX
CC   AAC97991 to AAC98763 encode the human colon cancer associated proteins,
CC   called human colon cancer antigens given in AAB53234 to AAB54006. The
CC   human colon cancer antigens can have cytostatic, cardioactive, muscular;
CC   neuroprotective, immunomodulatory, gynaecological, gastrointestinal,
CC   vulnerary, nephrotropic, antiinfective and antibacterial activities, and
CC   can be used in gene therapy. The colon cancer antigen polynucleotides,
CC   proteins and antibodies to the proteins are useful for the prevention,
CC   treatment and diagnosis of colon disorders, such as colon cancer. The
CC   polynucleotides may be used in diagnostics and research, such as for
CC   chromosome identification, and as hybridisation probes. The proteins
CC   may also be used to prevent diseases such as neural disorders, immune
CC   system disorders, muscular disorders, reproductive disorders,
CC   gastrointestinal disorders, wounds, renal disorders, infectious
CC   diseases, and cardiovascular disorders. AAC98764 to AAC99772 and
CC   AAB54007 represent sequences used in the exemplification of the present
CC   invention.
XX
SQ   Sequence   239 AA;


  Query Match             12.3%;  Score [illegible].5;  DB 21;  Length 239;
  Best Local Similarity   26.5%;  Pred. No. [illegible];
  Matches   57;  Conservative   40;  Mismatches  103;  Indels   15;  Gaps    7;

Qy       70 PQELAYL[illegible]LAMQKA[illegible]LNQ[illegible]NW[illegible]V[illegible]FVFRLFVVV 126
               |  ||   |  :  :| |     :     |      : : |    |   :      |     : : | |  : |
Db       29 dpala[illegible]qmkmlqyrdtaq[illegible]p[illegible]etpqn[illegible]yrqvdi 8[illegible]

Qy      127 DQPMERLYEELVERMEAMG      DWNPNV[illegible]KVD[illegible]HHLAALAANGNLV[illegible] 182
               :|  : : :   :      | : |      : | .  |     : : :          .         .      :
Db       84 yqtleevwdev  -kpavqqlrvkwde[illegible]tstpsaamkllspr 140

Qy      183 DFVGVK[illegible]AKE[illegible]GVLAGMAIDEQNM[illegible]KTKLLW 241
             ||| :    ||     |              :      |    |  .  :            |  |  : : | : :
Db      141 dtvdivlvkrvedqtissnathvelh[illegible]pk[illegible]k[illegible] 200

Qy      242 LLSIDLKGWLPKSIINQVLSQTQVDFANHLRKRLE 27[illegible]
                    ||  : :||: : : :        : :      |    :  :     : :
Db      201 tfhtdl[illegible]yl[illegible]qnvvdsttprsmtrtye[illegible]k[illegible]


RESULT   [illegible]5
AAB43727
ID   AAB43727 standard; Protein; [illegible] AA.
XX
AC   AAB43727;
XX
DT   08-FEB-2001  (first entry)
XX
DE   Human cancer associated protein sequence SEQ ID No:1172.
XX
KW   Human; cancer associated gene; cancer antigen; detection; cancer;
KW   diagnosis; cytostatic; proliferative; vulnerary; immunomodulator;
KW   antidiabetic; antiasthmatic; antirheumatic; antiarthritic; antiviral;
KW   antiinflammatory; [illegible]
KW   dermatological; neuroprotective; [illegible]
KW   vasotropic; [illegible]
KW   immune disorder; [illegible]
KW   allergic reaction; graft versus host [illegible]
KW   haemostatic; thrombolytic; [illegible]
KW   neurological disease; [illegible]
XX
OS   Homo sapiens.
XX
PN   WO2000[illegible] A1
XX
PD   21 SEP 2000
XX
PF   08 MAR 2000; [illegible]
XX
PR   12 MAR 1999; [illegible]
XX
PA   (HUMA-) HUMAN GENOME [illegible] INC.
XX
PI   Rosen CA, Ruben SM;
XX
DR   WPI; 2000-[illegible]
DR   N-PSDB; AAC[illegible]
XX
PT   Novel isolated [illegible]
PT   useful for treating [illegible]
XX
PS   Claim 11; Page [illegible]; [illegible]pp; [illegible]
XX
CC   AAC77[illegible] to AAC[illegible]118 [illegible]
CC   in AAB13998 to AAB44[illegible] The proteins can [illegible]
CC   tissues and cells. The [illegible]
CC   include cytostatic [illegible]
CC   antidiabetic; [illegible]
CC   antiinflammatory; antithyroid; [illegible]
CC   dermatological; neuroprotective; [illegible]
CC   nootropic; vasotropic; [illegible]
CC   polynucleotides and polypeptides can be used [illegible]
CC   amelioration of medical conditions [illegible]
CC   polynucleotides and polypeptides [illegible]
CC   the present invention may be used to treat [illegible]
CC   or inhibiting the proliferation [illegible]
CC   immune cells [illegible]
CC   disorders, allergic reactions, graft versus host [illegible]
CC   rejection, modulate [illegible]
CC   inflammation, cancers, cardiovascular [illegible]
CC   bacterial or viral infections. The peptides [illegible]
CC   agonists and antagonists may be [illegible]
CC   AAC784[illegible] and AAB11340 represent sequences [illegible]
CC   the present invention.
XX
SQ   Sequence [illegible] AA;

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches [illegible]; Conservative [illegible]; Mismatches [illegible]

Qy       69 YSDQFLAY[illegible]AMQ[illegible]
Db       52 [illegible]

Qy      129 QPMEELYE[illegible]MLAN[illegible]WNENVE[illegible]
Db      112 [illegible]


Search completed: Aug[illegible]
Job time: 40[illegible] sec
```

```
                       GenCore version 4.5
           Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on:         August 22, 2001, 14:25:30 ; Search time 64.49 Seconds
                                             (without alignments)
                                             336.637 Million cell updates/sec

Title:          US-08-659-254-18
Perfect score:  1470
Sequence:       1 MLLATFKLCAGSSYRHMRNM.............DFANHLRKRLESHPASEARC 285

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters:      219241

 mum DB seq length: 0
 mum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       PIR_68:*
               1:  pir1:*
               2:  pir2:*
               3:  pir3:*
               4:  pir4:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                              SUMMARIES
                  %
Result          Query
  No.   Score   Match Length DB  ID                      Description
------------------------------------------------------------------------------
    1    1464    99.6    285  2  I38248                  steroidogenic acut
    2    1299    88.4    285  2  JC4315                  steroidogenic acut
    3  1291.5    87.9    284  2  A55455                  steroidogenic acut
    4  1268.5    86.3    362  2  JC5386                  steroidogenic acut
    5     372    25.3    445  2  I38027                  MLN 64 protein - h
    6   182.5    12.4    241  2  T22[illegible]         [illegible]
    7     130     8.8    478  2  T16170                  hypothetical prote
    8   119.5     8.1    106  2  T46357                  hypothetical prote
    9   107.5     7.3    210  2  F69250                  [illegible]
   10   107.5     7.3   1281  2  G83405                  hypothetical prote
   11     104     7.1    988  2  A406[illegible]8       probable transposa
   12   101.5     6.9    239  2  C69251                  [illegible]
   13     101     6.9   1083  2  S54293                  regulator protein
   14     100     6.8    399  2  E71169                  hypothetical prote
   15    96.5     6.6    679  2  T05041                  hypothetical prote
   16      96     6.5    296  2  T23239                  hypothetical prote
   17      96     6.5    419  2  G750[illegible]        probable [illegible]
   18    94.5     6.4   3433  1  S28341                  utrophin - human
   19      94     6.4    322  2  T254[illegible]        hypothetical prote
   20      94     6.4    469  2  B70428                  general secretion
   21      94     6.4    753  1  D726[illegible]        probable [illegible]
   22      93     6.3    354  2  E711[illegible]3       probable operon pr
   23      [illegible] 6.2 15[illegible] 2 S49414        [illegible]
   24      90     6.1   1188  2  G839[illegible]        chromosome segrega
   25      89     6.1    607  2  H84514                  hypothetical prote
   26      89     6.1    632  2  S38012                  probable [illegible]
   27      89     6.1   1019  2  I152[illegible]        hypothetical prote
   28      89     6.1   1529  2  T20986                  hypothetical prote
   29    88.5     6.0    422  2  S244[illegible]        hypothetical prote
   30    88.5     6.0    863  2  B69391                  signal-transducing
   31    88.5     6.[illegible] 1055 2 F64114           type I site-specif
   32      88     6.[illegible] 394 2 S370[illegible]1  4-hydroxybenzoate
   33      88     [illegible] 949 2 T18541               hypothetical prote
   34    87.5     6.0   1343  1  H64073                  DNA-directed RNA p
   35      87     5.9    425  2  D72345                  trigger factor [illegible]
   36      87     5.9    643  2  I59412                  sepA protein - [illegible]
   37      87     5.9    675  2  E86646                  escV (imported)
   38      87     5.9   1033  2  I487[illegible]        Smcx protein ([illegible]
   39    86.5     5.9    423  1  I40382                  isocitrate dehydro
   40    86.5     5.9   1225  1  B64234                  hypothetical prote
   41      86     5.9    218  2  A69[illegible]2        conserved hypothet
   42      86     5.9    303  2  S01295                  intermediate filam
   43      86     5.9    978  2  A70387                  conserved hypothet
   44      86     5.9   1389  2  T30834                  RNA polymerase bet
   45      86     5.9   1896  2  T01499                  hypothetical prote

                              ALIGNMENTS

RESULT    1
I38248
steroidogenic acute regulatory protein - human
C;Species: Homo sapiens (man)
C;Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 05-Nov-1999
C;Accession: I38248; I38896
R;Sugawara, T.; Lin, D.; Holt, J.A.; Martin, K.O.; Javitt, N.B.; Miller, W.L.; Straus
Biochemistry 34, 12506-12512, 1995
A;Title: Structure of the human steroidogenic acute regulatory protein (StAR) gene: S
A;Reference number: I38248; MUID:960[illegible]
A;Accession: I38248
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-285 <RES>
A;Cross-references: EMBL:U29105; NID:g[illegible]; PIDN:AA[illegible]234.1; PID:g[illegible]
R;Sugawara, T.; Holt, J.A.; Driscoll, D.; Strauss III, J.F.; Lin, D.; Miller, W.L.; P
Proc. Natl. Acad. Sci. U.S.A. 92, 4778-4782, 1995
A;Title: Human steroidogenic acute regulatory protein: functional activity in COS-1 c
A;Reference number: I38896; MUID:95281[illegible]
A;Accession: I38896
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-285 <[illegible]>
A;Cross-references: EMBL:U17280; NID:g[illegible]; PIDN:AA[illegible]; PID:g[illegible]
C;Genetics:
A;Gene: StAR
A;Cross-references: GDB:STAR; GDB:[illegible]; OMIM:600617
A;Map position: 8p11.2-8p11.2
A;Introns: [illegible]

  Query Match             99.6%;  Score 1464;  DB 2;  Length 285;
  Best Local Similarity   99.6%;  Pred. No. [illegible];
  Matches  284;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 MLLATFKLCAGSSYRHMRNMKGLRQQAVMAISQELNRRALGGPTPSTWINQVRRRSSLLG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MLLATFKLCAGSSYRHMRNMKGLRQQAVMAISQELNRRALGGPTPSTWINQVRRRSSLLG 60

Qy         61 SRLEETLYSDQELAYLQQGEEAMQKALGILSNQEGWKKESQQDNGDKVMSKVVPDVGKVF 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 SRLEETLYSDQELAYLQQGEEAMQKALGILSNQEGWKKESQQDNGDKVMSKVVPDVGKVF 120

Qy        121 RLEVVVDQPMERLYEELVERMEAMGEWNPNVKEIKVLQKIGKDTFITHELAAEAAGNLVG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 RLEVVVDQPMERLYEELVERMEAMGEWNPNVKEIKVLQKIGKDTFITHELAAEAAGNLVG 180

Qy        181 PRDFVSVRCAKRRGSTCVLAGMATDFGNMPEQKGVIRAEHGPTCMVLHPLAGSPSKTKLT 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PRDFVSVRCAKRRGSTCVLAGMATDFGNMPEQKGVIRAEHGPTCMVLHPLAGSPSKTKLT 240

Qy        241 WLLSIDLKGWLPKSIINQVLSQTQVDFANHLRKRLESHPASEARC 285
```

```
Db    241 WEESIDEKGWEEKS [illegible] NQVE [illegible] VEEANHP [illegible]


RESULT    2
JC4315
steroidogenic acute regulatory protein [illegible]
C;Species: Bos primigenius taurus (cattle)
C;Date: 29 Nov-1996 #sequence_revision [illegible]
C;Accession: JC4315
R;Hartung, [illegible]; Rust, W.; [illegible], M. [illegible]
Biochem. Biophys. Res. Commun. 215 [illegible]
A;Title: Molecular [illegible] and [illegible]
A;Reference number: [illegible] MUID [illegible]
A;Accession: JC4315
A;Molecule type: mRNA
A;Residues: 1-285 [illegible]
A;Comment: This protein is an acute [illegible]
[illegible]netics.
A;Gene: STAR
F;226-264/Region: [illegible]


  Query Match             [illegible]
  Best Local Similarity   [illegible]  Pred. No. [illegible]
  Matches  [illegible];  Conservative  [illegible];  Mismatches  [illegible]

Qy      1 MLLATFKLCAG [illegible] YRHMRNMKGLRQQAVMAI [illegible] NLRAL [illegible]

Db      1 MLLATFKLCAG [illegible] YRHVSMKGLRQQAVIA [illegible] NRSA [illegible]

Qy     61 SREEEILYS [illegible] AY [illegible] EAMQKA [illegible] NQE [illegible] WKKE [illegible]

Db     61 SQLEDPLYSDQELAH [illegible] GELAMQKA [illegible] KDQE [illegible] WKKE [illegible]

Qy    121 RLEVVVDQPMELLYEE [illegible] VERMEAMGEWNPNVKEIKVLQK [illegible]

Db    121 RLEVVVDQPMERLYEE [illegible] VERMEAMGEWNPNVKEIKVLQK [illegible]

Qy    181 PREEVSVE [illegible] AK [illegible] ISG [illegible] VLA [illegible] MA [illegible] NMPE [illegible] K [illegible]

Db    181 PRGEVSVE [illegible] K [illegible] GSM [illegible] VLAGMATLYEEMPS [illegible] K [illegible]

Qy    241 WEESIDEKGWEEKSEENQVEQ [illegible] VEEANHPKKPE [illegible] A [illegible]

Db    241 WEESIDEKGWEEKEE [illegible] VE [illegible] VEEANH [illegible]


RESULT    3
A55455
steroidogenic acute regulatory protein [illegible]
C;Species: Mus musculus (house mouse)
C;Date: 19 Feb 1997 #sequence_revision [illegible]
C;Accession: A55455
R;Clark, B.J.; Wells, [illegible]
J. Biol. Chem. 269 [illegible] 34 [illegible]
A;Title: The purification [illegible] expression [illegible]
ry protein (StAR)
A;Reference number: A55455 MUID [illegible]
A;Accession: A55455
A;Status: preliminary [illegible] compared with [illegible]
A;Molecule type: mRNA
A;Residues: 1-284 [illegible]
A;Cross-references: [illegible]
C;Keywords: mitochondri[illegible]
F;1-47/Domain: transit peptide (mitochondrion) #status predicted [illegible]


  Query Match             [illegible]
  Best Local Similarity   [illegible]  Pred. No. [illegible]
  Matches  248;  Conservative  [illegible];  Mismatches  [illegible]

Qy      1 MLLATFKLCAG [illegible] YRHM [illegible] NMK [illegible] QAVMA [illegible]
```

```
Db      [illegible] WMQVRRRS [illegible]

Qy      [illegible] AY [illegible] EAMQKA [illegible] NQE [illegible] WKKE [illegible] KVMCKVVC [illegible] VGKVE [illegible]

Db      [illegible] YSDQELAY [illegible] GEVAMQKA [illegible] NNQE [illegible] WKK [illegible] KVMCKVVPVGKVE [illegible]

Qy    121 [illegible] VVVDQPME [illegible] VERMEAMGEWNPNVKEIKV [illegible] LQ [illegible] THIL [illegible] AAIAA [illegible] NLV [illegible]

Db    [illegible] VVVDQPME [illegible] AMGEWNPNVKEIKV [illegible] VTHHLAAAAA [illegible] NLV [illegible]

Qy    [illegible] VE [illegible] AK [illegible] MA [illegible] NMP [illegible] K [illegible] MVIHL [illegible] KIK [illegible] 234

Db    [illegible] VSVE [illegible] VLA [illegible] MA [illegible] EMPE [illegible] MVIHP [illegible] KIK [illegible]

Qy    [illegible] WE [illegible] KGWEEK [illegible] NQVE [illegible] VEEANHPKKPE [illegible] A [illegible] 285

Db    [illegible] ANG [illegible] A [illegible]


RESULT    4
[illegible]
steroidogenic acute regulatory protein [illegible]
C;Species: Rattus [illegible] (Norway rat)
C;Date: [illegible] #sequence_revision [illegible] #text_change [illegible] May [illegible]
C;Accession: [illegible]
R;Lee, [illegible]
[illegible]
A;Title: [illegible] acute regulatory protein [illegible]
A;Reference number: [illegible] MUID [illegible]
A;Accession: [illegible]
A;Molecule type: mRNA
A;Residues: [illegible]
A;Cross-references: [illegible]
C;Comment: This protein is a key regulator [illegible] steroidogenesis in acute response [illegible]


  Query Match             [illegible]  Score [illegible]  DB [illegible]  Length [illegible]
  Best Local Similarity   [illegible]  Pred. No. [illegible]
  Matches  [illegible];  Conservative  [illegible];  Mismatches  [illegible];  Indels  [illegible];  Gaps  [illegible]

Qy      [illegible] NMKGLRQQAVMAISQ [illegible] NLRAL [illegible] WIN [illegible] VRRRSSLLSSR [illegible]

Db      [illegible] NMKGLRHQAVIA [illegible] NLRAL [illegible] WMQVRRRGSLL [illegible] 

Qy      [illegible] AM [illegible] KAL [illegible] NQE [illegible] WKKE [illegible] KVMCKVVP [illegible] VGKV [illegible] EV 124

Db      [illegible] AL [illegible] YS [illegible] EAMQKALG [illegible] NNQEGWKKE [illegible] VISKVVPDVGKVFRLEV [illegible]

Qy      [illegible] VVD [illegible] MEAMGEWNPNVKEIKV [illegible] LAAIAA [illegible] NLV [illegible] 184

Db      [illegible] MEAMGEWNPNVKEIKV [illegible] LAAAAA [illegible] NLV [illegible]

Qy      [illegible] VSVE [illegible] AKR [illegible] A [illegible] MA [illegible] NMP [illegible] K [illegible] MVI [illegible] LA [illegible] KIK [illegible] WE [illegible] 244

Db      [illegible] VSVE [illegible] KR [illegible] A [illegible] MA [illegible] MP [illegible] K [illegible] MVI [illegible] LA [illegible] KIK [illegible] WE [illegible]

Qy      [illegible] KGWEEK [illegible] NQVE [illegible] VEEANPKKPE [illegible] A [illegible]

Db      [illegible] EA [illegible] A [illegible]


RESULT    [illegible]
[illegible]
[illegible]
C;Species: [illegible]
C;Date: [illegible] Nov [illegible] #sequence_revision [illegible] Nov [illegible] #text_change [illegible] Nov 1999
C;Accession: [illegible]
R;[illegible]
[illegible]
A;Title: [illegible] and [illegible] overexpressed [illegible]
A;Reference number: [illegible] MUID [illegible]
A;Accession: [illegible]
```

```
                    GenCore version 4.5
          Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on:         August 22, 2001, 14:30:1[illegible] ; Search time [illegible].29 Seconds
                                          (without alignments)
                                          106.136 Million cell updates/sec

Title:          US-08-659-254-18
Perfect score:  1470
Sequence:       1 MLLATFKLCAGSSYRHMRNM ......    ...[illegible] 285

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       197339 seqs, 20590346 residues

Total number of hits satisfying chosen parameters:      197339

  [illegible]imum DB seq length: 0
  [illegible]imum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       Issued_Patents_AA:*
               1:  [illegible]_COMB.pep:*
               2:  [illegible]_COMB.pep:*
               3:  [illegible]_COMB.pep:*
               4:  /cgn2_6/ptodata/2/iaa/[illegible]_COMB.pep:*
               5:  [illegible]_COMB.pep:*
               6:  [illegible].pep:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                                  SUMMARIES
                  %
Result          Query
  No.   Score   Match Length DB  ID                       Description
------------------------------------------------------------------------------
   1    1470    100.0   285  4  US-08-659-254-18         Sequence 18, Appli
   2    1464     99.6   285  1  US-08-410-[illegible]-2          Sequence 2, Appli
   3  1291.5     87.9   284  2  US-08-538-[illegible]-2          Sequence 2, Appli
   4  1258.5     85.6   276  4  US-08-659-254-2          Sequence 2, Appli
   5     372     25.3   445  2  US-08-691-[illegible]-6          Sequence 6, Appli
   6     104      7.1   988  2  US-08-286-[illegible]A-19        Sequence 19, Appl
   7     104      7.1   988  3  US-08-980-[illegible]-19         Sequence 19, Appl
   8    98.5      6.7   291  4  US-09-328-[illegible]-2          Sequence 2, Appli
   9    90.5      6.2   494  4  US-09-198-[illegible]-2          Sequence 2, Appli
  10      89      6.1    19  2  US-08-[illegible]-5              Sequence 5, Appli
  11      89      6.1    19  4  US-08-659-254-5          Sequence 5, Appli
  12      83      5.6   402  4  US-09-164-[illegible]-4          Sequence 4, Appli
  13      83      5.6   402  4  US-09-414-[illegible]-4          Sequence 4, Appli
  14      82      5.6   600  5  PCT-US9[illegible]-2             Sequence 2, Appli
  15      81      5.5   529  4  US-09-4[illegible]-2             Sequence 2, Appli
  16      81      5.5   529  4  US-09-414-[illegible]-2          Sequence 2, Appli
  17      80      5.4    14  2  US-08-538-[illegible]-3          Sequence 3, Appli
  18      80      5.4    14  4  US-08-659-254-3          Sequence 3, Appli
  19    79.5      5.4   942  3  US-09-07[illegible]-1            Sequence 1, Appli
  20    79.5      5.4   942  4  US-09-388-7[illegible]-1         Sequence 1, Appli
  21      79      5.4   974  2  US-08-868-7[illegible]-6         Sequence 6, Appli
  22    78.5      5.3   786  3  US-09-01[illegible]-11           Sequence 11, Appl
  23    77.5      5.3   255  1  US-08-459-2[illegible]-4         Sequence 4, Appli
  24    77.5      5.3   255  1  US-08-459-2[illegible]-4         Sequence 4, Appli
  25    77.5      5.3   572  2  US-08-74[illegible]-4-[illegible]       Sequence 3, Appli
  26    77.5      5.3   572  4  US-09-14[illegible]               Sequence 6, Appli
  27    76.5      5.2   222  2  US-08-79[illegible]-2            Sequence 2, Appli
  28    76.5      5.2   222  4  US-09-3[illegible]               Sequence [illegible], Appli
  29    76.5      5.2   573  2  US-08-71[illegible]              Sequence 4, Appli
  30    76.5      5.2   573  4  US-09-1[illegible]               Sequence 7, Appli
  31      76      5.2   310  3  US-08-9[illegible]-4             Sequence 4, Appli
  32      76      5.2   335  1  US-08-3[illegible]               Sequence 2, Appli
  33      76      5.2   335  2  US-08-4[illegible]               Sequence 2, Appli
  34      75      5.1   185  2  US-08-[illegible]-12             Sequence 12, Appl
  35      75      5.1   872  1  US-08-7[illegible]               Sequence 2, Appli
  36    74.5      5.1   793  3  US-09-0[illegible]               Sequence [illegible], Appl
  37      74      5.0   699  3  US-08-8[illegible]-84A-[illegible]          Sequence 52, Appl
  38      74      5.0   699  4  US-08-97[illegible]-54A-[illegible]         Sequence 188, App
  39      74      5.0   719  3  US-08-8[illegible]-84A-[illegible]          Sequence 7, Appli
  40      74      5.0   719  4  US-08-97[illegible]-54A-219      Sequence 219, App
  41      74      5.0   727  4  US-08-[illegible]                Sequence 4, Appli
  42      74      5.0   760  4  US-08-928-[illegible]            Sequence 2[illegible], Appl
  43      74      5.0  1039  6  5196511-2                Patent No. 5196511
  44      74      5.0  2890  4  US-09-41[illegible]-6[illegible]           Sequence 6[illegible], Appl
  45      74      5.0  3798  3  US-09-[illegible]-6              Sequence 6, Appli

                                 ALIGNMENTS

RESULT 1
US-08-659-254-18
; Sequence 18, Application US/08659254
; Patent No. 6194[illegible]
;  GENERAL INFORMATION:
;    APPLICANT:  Stocco, Douglas M.
;    APPLICANT:  Clark, Dr. Barbara J.
;    TITLE OF INVENTION:  COMPOSITIONS AN[illegible]
;    TITLE OF INVENTION:  OF STEROIDOGENESIS
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  [illegible]ump, Strauss, Hauer & Feld, L.L.P.
;      STREET:  [illegible] Avenue
;      CITY:  Austin
;      STATE:  TX
;      COUNTRY:  U.S.A.
;      ZIP:  78701
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/659,254
;      FILING DATE:  07-JUN-1996
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/538,9[illegible]
;      FILING DATE:  04-OCT-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Mayfield, Denise L.
;      REGISTRATION NUMBER:  33,732
;      REFERENCE/DOCKET NUMBER:  43375.0006
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  512/499-6200
;      TELEFAX:  512/499-6290
;  INFORMATION FOR SEQ ID NO:  18:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  285 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
US-08-659-254-18

  Query Match             100.0%;  Score 1470;  DB 4;  Length 285;
  Best Local Similarity   100.0%;  Pred. No. [illegible];
  Matches  285;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 MLLATFKLC[illegible]VPRSQ[illegible] 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MELATEKL[illegible]YR[illegible]ENMK[illegible]LAVMA[illegible]NERA[illegible]

Qy         61 [illegible]ELLY[illegible]AY[illegible]AMQKA[illegible]N[illegible]WKKL[illegible]N[illegible]
              [illegible]
Db         61 [illegible]ELY[illegible]AY[illegible]AMQKA[illegible]N[illegible]WKKL[illegible]N[illegible]

Qy        121 [illegible]EVVV[illegible]MEE[illegible]YEE[illegible]VE[illegible]MEAM[illegible]WN[illegible]NVK[illegible]
              [illegible]
Db        121 [illegible]EVVV[illegible]ME[illegible]YEE[illegible]VE[illegible]MEAM[illegible]WN[illegible]NVK[illegible]

Qy        181 [illegible]VSV[illegible]A[illegible]VEA[illegible]MA[illegible]NM[illegible]
              [illegible]
Db        181 [illegible]VSV[illegible]AK[illegible]VEA[illegible]MA[illegible]NM[illegible]

Qy        241 WE[illegible]E[illegible]W[illegible]V[illegible]ANG[illegible]
              [illegible]
Db        241 WE[illegible]K[illegible]W[illegible]V[illegible]ANG[illegible]

RESULT    2
US-08-41[illegible]-2
; Sequence 2, Application [illegible]
; Patent No. [illegible]
;  GENERAL INFORMATION:
;    APPLICANT: [illegible]
;    APPLICANT: [illegible]
;    APPLICANT: [illegible] Jerome [illegible]
;    TITLE OF INVENTION: [illegible]
;    TITLE OF INVENTION: [illegible] WITH [illegible]
;    NUMBER OF SEQUENCES: [illegible]
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE: [illegible]
;      STREET: [illegible] Alto Square
;      CITY: Palo Alto
;      STATE: CA
;      COUNTRY: US
;      ZIP: [illegible]
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE: Floppy disk
;      COMPUTER: IBM PC compatible
;      OPERATING SYSTEM: PC-DOS/MS-DOS
;      SOFTWARE: PatentIn Release #[illegible], Version #[illegible]
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER: [illegible]
;      FILING DATE: [illegible] MAR [illegible]
;      CLASSIFICATION: 43[illegible]
;    ATTORNEY/AGENT INFORMATION:
;      NAME: Neeley, Richard L.
;      REGISTRATION NUMBER: [illegible]
;      REFERENCE/DOCKET NUMBER: [illegible]
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE: 41[illegible]
;      TELEFAX: 41[illegible]
;      TELEX: 380816[illegible]
;  INFORMATION FOR SEQ ID NO: [illegible]
;    SEQUENCE CHARACTERISTICS:
;      LENGTH: 28[illegible] amino acids
;      TYPE: amino acid
;      TOPOLOGY: [illegible]
;    MOLECULE TYPE: protein
US-08-41[illegible]

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches  28[illegible]; Conservative [illegible]; Mismatches [illegible]

Qy          1 MELATEKLA[illegible]RHM[illegible]NMK[illegible]LAVMA[illegible]
              [illegible]
Db          1 MELATEKLA[illegible]RHM[illegible]NMK[illegible]LAVMA[illegible]

Qy         61 SRLEEELLYSE[illegible]AY[illegible]AMQKA[illegible]N[illegible]WKK[illegible]
              [illegible]
```

```
Qy        [illegible] KVM[illegible]VVE[illegible]

Qy        [illegible]VVV[illegible]ME[illegible]VE[illegible]MEAM[illegible]WN[illegible]NVK[illegible]HHE[illegible]AAEAA[illegible]
              [illegible]
Db        [illegible]VVV[illegible]ME[illegible]VE[illegible]MEAM[illegible]WN[illegible]NVK[illegible]HHE[illegible]AAEAA[illegible]

Qy        [illegible]NME[illegible]MV[illegible]
              [illegible]
Db        [illegible]NME[illegible]MVEH[illegible]

Qy        [illegible]W[illegible]K[illegible]N[illegible]ANG[illegible]
              [illegible]
Db        [illegible]W[illegible]K[illegible]N[illegible]ANG[illegible]

RESULT [illegible]
[illegible]
; Sequence [illegible], Application [illegible]
; Patent No. [illegible]
;  GENERAL INFORMATION:
;    APPLICANT: [illegible]
;    APPLICANT: [illegible]
;    TITLE OF INVENTION: [illegible] COMPOSITIONS AND METHODS [illegible]
;    TITLE OF INVENTION: [illegible]
;    NUMBER OF SEQUENCES: [illegible]
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE: [illegible]
;      STREET: [illegible]
;      CITY: [illegible]
;      STATE: [illegible]
;      COUNTRY: [illegible]
;      ZIP: [illegible]
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE: [illegible]
;      COMPUTER: IBM [illegible]
;      OPERATING SYSTEM: [illegible]
;      SOFTWARE: [illegible]
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER: [illegible]
;      FILING DATE: [illegible]
;      CLASSIFICATION: [illegible]
;    ATTORNEY/AGENT INFORMATION:
;      NAME: [illegible]
;      REGISTRATION NUMBER: [illegible]
;      REFERENCE/DOCKET NUMBER: [illegible]
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE: [illegible]
;      TELEFAX: [illegible]
;  INFORMATION FOR SEQ ID NO: [illegible]
;    SEQUENCE CHARACTERISTICS:
;      LENGTH: [illegible]
;      TYPE: [illegible]
;      STRANDEDNESS: [illegible]
;      TOPOLOGY: [illegible]
;    MOLECULE TYPE: [illegible]
[illegible]

  Query Match [illegible]
  Best Local Similarity [illegible]
  Matches [illegible]; Mismatches [illegible]

Qy        [illegible] NMK[illegible]LAVMA[illegible]NERA[illegible]
              [illegible]
Db        [illegible] NMK[illegible]LAVLA[illegible]

Qy        [illegible]AMQKA[illegible]N[illegible]WKK[illegible]KVM[illegible]KVVE[illegible]
              [illegible]
Db        [illegible]AMQKA[illegible]NN[illegible]WKK[illegible]KMVE[illegible]

Qy        [illegible]EEVVV[illegible]ME[illegible]VEEMEAMGWN[illegible]NVK[illegible]HHEAAEAAGN[illegible]
              [illegible]
Db        [illegible]VVV[illegible]ME[illegible]MEAM[illegible]WN[illegible]NVK[illegible]HHE[illegible]AAAAA[illegible]
```

```
Qy        181 PPDFVSVRCAKRRGSTCVLAGMATDFGNMPEQKGVIKAEHGPTCMVLHPLAGSPSKIKLT 240
              |:||||||| |||||||||||||||| || ||||   |   ||||||||||||||||:|:|:|:|:
Db        180 PRDFVSVRCTKRRGSTCVLAGMATHFGEMPEQSGVIKAEHGPTCMVLHPLAGSPSKIKLT 239

Qy        241 WLLSIDLKGWLPKSIINQVLSQTQVDFANHLRKKLESHPASEARC 285
              |||||||||||||:|||||||||:::|||||:::|: |||||:|
Db        240 WLLSIDLKGWLPKTIINQVLSQTQIEFANHLRKKLEASPASEAQC 284


RESULT    4
US-08-659-254-2
; Sequence 2, Application US/08659254
; Patent No. 6194555
;  GENERAL INFORMATION:
;    APPLICANT:  Stocco, Douglas M.
;    APPLICANT:  Clark, Dr. Barbara J.
;    TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR REGULATION
;    TITLE OF INVENTION:  OF STEROIDOGENESIS
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Akin, Gump, Strauss, Hauer & Feld, L.L.P.
;      STREET:  1900 Frost Bank Plaza, 816 Congress Avenue
;      CITY:  Austin
;      STATE:  TX
;      COUNTRY:  U.S.A.
;      ZIP:  78701
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/659,254
;      FILING DATE:  07-JUN-1996
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/538,060
;      FILING DATE:  04-OCT-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Mayfield, Denise L.
;      REGISTRATION NUMBER:  33,732
;      REFERENCE/DOCKET NUMBER:  43375.0006
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  512/499-6200
;      TELEFAX:  512/499-6290
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  276 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-659-254-2

  Query Match             85.6%;  Score 1275.5;  DB 4;  Length 276;
  Best Local Similarity   87.4%;  Pred. No. 2.7e-130;
  Matches  242;  Conservative   18;  Mismatches   16;  Indels    0;  Gaps    1;

Qy          1 MLLATFKLCAGSSYRHMRNMKGLRQQAVMAIGQELNRRALGGPTPSIWINQVRRRSSLLG 60
              | |||||||||||||||||||||||| |||:|| | || ||||  :|  |: |||:|||||
Db          1 MFLATFKLCAGSSYRHMRNMKGLRHQAVLAIGQELNWRALGDSSPG WMGQVRRRSSLLG 59

Qy         61 SRLEETLYSDQELAYLQQGEEAMQKALGILSNQEGWKKESQQDNGLKVMSKVVPDVGKVF 120
              |:|| ||||||||:|:|||| ||||||||| |||||||||||:|||  |:|:|:|||||||
Db         60 SQLEATLYSDQELSYIQQGEVAMQKALGILNNQEGWKKESQQDNGDKVMSKMVPDVGKVF 119

Qy        121 RLEVVVDQPMERLYEELVERMEAMGEWNPNVKEIKVLQKIGKDTFITHELAAEAAGNLVG 180
              ||||||||||:|||||||:||||||| || |:| |||:||||| |||:||| ||||||
Db        120 RLEVVVDQPMDRLYEELVDRMEAMGEWNPNVKEIKVLQRIGKDTVITHELAAAAAGNLVG 179

Qy        181 PRDFVSVRCAKRRGSTCVLAGMATDFGNMPEQKGVIKAEHGPTCMVLHPLAGSPSKIKLT 240
```

```
              |||||||||  |||||||||||||| || ||       :   :||||| ||||||:|:||
Db        180 PRDFVSVRCTKRRGSTCVLAGMATHFGEMPEQSGVIKAEHGPTCMVLHPLAGSPSKIKL  239

Qy        241 WLLSIDLKGWLPKSIINQVLSQTQVDFANHLRKKLES 277
              |||||||||||||:||||||||||::| || || | :
Db        240 WLLSIDLKGWLPKTIINQVLSQTQIEFANHLRKKLEA 276


RESULT    5
US-08-691-814B-6
; Sequence 6, Application US/08691814B
; Patent No. 5981218
;  GENERAL INFORMATION:
;    APPLICANT:  Rio, Marie-Christine
;    APPLICANT:  Tomasetto, Catherine
;    APPLICANT:  Basset, Paul
;    APPLICANT:  Byrne, Jennifer
;    TITLE OF INVENTION:  Isolated Nucleic Acid Molecules Useful
;    TITLE OF INVENTION:  as Leukemia Markers and in Breast Cancer Prognosis
;    NUMBER OF SEQUENCES:  124
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sterne, Kessler, Goldstein & Fox P.L.L.C.
;      STREET:  1100 New York Ave, NW, Suite 600
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20005-3934
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.3
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/691,814
;      FILING DATE:  31-JUL-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/002,83
;      FILING DATE:  09-AUG-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Steffe, Eric K.
;      REGISTRATION NUMBER:  36,688
;      REFERENCE/DOCKET NUMBER:  1383.009-001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-371-2600
;      TELEFAX:  202-371-2543
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  445 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-691-814B-6

  Query Match             25.3%;  Score 372;  DB 2;  Length 445;
  Best Local Similarity   37.5%;  Pred. No. 2e-32;
  Matches   78;  Conservative   45;  Mismatches   85;  Indels    0;  Gaps    0;

Qy         68 YSDQELAYLQQGEEAMQKALGILSNQEGWKKESQQDNGLKVMSKVVPDVGKVFRLEVVVD 127
              :| ||  |::||:||      ||: :|  | |   :  ,  |   ||   || | |:  :
Db        231 FSAQERESIPQCKEATAVVDQILAQEENWKEEKNNEYSDIVYSIEVPEHGKIFILKEEP 290

Qy        128 QPMERLYEELVERMEAMGEWNPNVKEIEVIRKICKEIFITHEIAAEAAGNLVGPRDFVSV 187
                | | :|:|::: : | | || |    :||::   | |:::::| ||| :  || ||:|
Db        291 CPAELVYQLVIQAERMVLWNKTVEACQIFLRVEENTLISYIKSAGAAGGVVSPREFVNV 350

Qy        188 RCAKRRGSTCVLAGMATDFGNMPEQKGVIKAEHGPTCMVLHPLAGSPSKIKLIWLLSIDL 247
               |   :||        :  |        |  :  |:||   ::    :|     :|:|: ||
Db        351 RRIERRRDGSCIATSHSAKPHSHKYVRGENGPGGIVLKSASNPRVCIFVWILNIDL 410

Qy        248 KGWLPKSIINQVLSQTQVDFANHLRKKL 275
              || ||: :|:| |: |  :|| |||:|:
```

```
Db      411 KGRLPPYL[illegible]SLAAIMEEAPHER[illegible] [illegible]

RESULT    [illegible]
US-08-286-819A-19
; Sequence 19, Application US/08286819A
; Patent No. 587[illegible]
;  GENERAL INFORMATION:
;    APPLICANT:  ARTHUR, MICHEL
;    APPLICANT:  DUTKA MALEN, SYLVIE
;    APPLICANT:  MOLINAS, CATHERINE
;    APPLICANT:  COURVALIN, PATRICE
;    TITLE OF INVENTION:  POLYPEPTIDES IMPLICATED IN THE
;    TITLE OF INVENTION:  EXPRESSION OF RESISTANCE [illegible] IN PARTICULAR
;    TITLE OF INVENTION:  IN GRAM POSITIVE BACTERIA [illegible]
;    TITLE OF INVENTION:  THESE POLYPEPTIDES AND [illegible]
;    NUMBER OF SEQUENCES:  [illegible]4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  OBLON, SPIVAK, MCCLELLAND, MAIER & N[illegible]
;      ADDRESSEE:  P.C.
;      STREET:  17[illegible] S. Jefferson Davis Highway, [illegible]
;      CITY:  Arlington
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  2220[illegible]
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.[illegible]
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/286,819A
;      FILING DATE:  [illegible]-AUG-1994
;      CLASSIFICATION:  4[illegible]
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/1[illegible]4,[illegible]82
;      FILING DATE:  28-DEC-199[illegible]
;      CLASSIFICATION:  4[illegible]
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/9[illegible] 146
;      FILING DATE:  [illegible] 1992
;      CLASSIFICATION:  4[illegible]
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/FR/9[illegible]/00[illegible]
;      FILING DATE:  2[illegible]-OCT-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  FR 90[illegible]
;      FILING DATE:  [illegible]-OCT-1990
;      CLASSIFICATION:  43
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Oblon, N[illegible]
;      REGISTRATION NUMBER:  24,618
;      REFERENCE/DOCKET NUMBER:  [illegible]
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 4[illegible]
;      TELEFAX:  (703) 413[illegible]
;      TELEX:  248855 [illegible]
;  INFORMATION FOR SEQ ID NO: 1[illegible]:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9[illegible] amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-286-819A-19

  Query Match   [illegible];  Score [illegible];  DB [illegible];  Length [illegible];
  Best Local Similarity   [illegible];  Pred. No. [illegible];
  Matches [illegible];  Conservative [illegible];  Mismatches [illegible];  Indels [illegible];  Gaps [illegible];

Qy    PTPSTWINQVE[illegible]
Db    [illegible]
Qy    SPETFLKTIEKLEY[illegible]MDLETV[illegible]

Qy     [illegible]   [illegible]AYI[illegible]AMLKA     [illegible]NGDKVMSKVV [illegible]
Db     [illegible]YL[illegible]KAFF[illegible]SKGKFA[illegible]NGKKLNPKV[illegible]
Qy     [illegible] VPEL[illegible]VVV[illegible]YPELVLEMEAM[illegible]EWNPNVK [illegible]AAFAA[illegible]
Db     [illegible]  LAI KAFEKKI   VFKVLESVTEWN        [illegible]
Qy     [illegible]
Db     [illegible] AFLALY       [illegible]KAN 41[illegible]
Qy     [illegible]NGVL[illegible]ANHLKK[illegible]
Db     [illegible]

RESULT    [illegible]
US-[illegible]
; Sequence [illegible], Application [illegible]
; Patent No. [illegible]
;  GENERAL INFORMATION:
;    APPLICANT:  ARTHUR, MICHEL
;    APPLICANT:  DUTKA MALEN, SYLVIE
;    APPLICANT:  MOLINAS, CATHERINE
;    APPLICANT:  COURVALIN, PATRICE
;    TITLE OF INVENTION:  POLYPEPTIDES IMPLICATED IN THE
;    TITLE OF INVENTION:  EXPRESSION OF RESISTANCE TO GLYCOPEPTIDES, IN PARTICULAR
;    TITLE OF INVENTION:  IN GRAM POSITIVE BACTERIA, NUCLEOTIDE SEQUENCE CODING FOR
;    TITLE OF INVENTION:  THESE POLYPEPTIDES AND USE FOR DIAGNOSIS
;    NUMBER OF SEQUENCES:  [illegible]4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT
;      ADDRESSEE:  P.C.
;      STREET:  1755 S. Jefferson Davis Highway, [illegible]
;      CITY:  Arlington
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  [illegible]
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #[illegible], Version #[illegible]
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/[illegible]
;      FILING DATE:
;      CLASSIFICATION:  [illegible]
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/286,819
;      FILING DATE:  [illegible] AUG 1994
;      APPLICATION NUMBER:  US 08/174,682
;      FILING DATE:  28 [illegible]
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/917,146
;      FILING DATE:  [illegible]
;    PRIOR APPLICATION DATA
;      APPLICATION NUMBER:  [illegible]
;      FILING DATE:  [illegible]
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  [illegible]
;      FILING DATE:  [illegible]
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  [illegible]
;      REGISTRATION NUMBER:  [illegible]
;      REFERENCE/DOCKET NUMBER:  [illegible]
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) [illegible]
;      TELEFAX:  (703) 4[illegible]
;      TELEX:  [illegible]
;  INFORMATION FOR SEQ ID NO: [illegible]:
;    SEQUENCE CHARACTERISTICS:
```

```
;       LENGTH:  988 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-980 357-19

  Query Match               7.1%;  Score 104  [illegible]  Length 988;
  Best Local Similarity   20.7%;  Pred. No. [illegible];
  Matches   58,  Conservative   50;  Mismatches   80;  Indels   92;  Gaps   12;

Qy       41 GGPTPSTWINQVRR--------------------[illegible] 72
             | |:| |::   : |                        : |  ||||  |    :  | ||
Db      221 GHPSPETFLKIIFRLPYIRGMDLETVQISHL[illegible] 2[illegible]

Qy       73 ----LAYLQQGEEAMQKA -------LGI[illegible] 118
                :  ||   :|    ||          | :||       ::| |: || |:   |:    ::|
Db      281 SILTIYLLQLTQELIDKAFEIHDRQILSL[illegible] [illegible]

Qy      119 VFRLEVVVDQPMERLYEELVERMEAMCEW[illegible] 1[illegible]
                :  ::        :|   ::  :  :|::  ||                 ||::     :  |  |    |
Db      340 ----QALIKAREEKL--DVFKVLESVIEWN-[illegible]EVS---SVEEAQEL 37[illegible]

        179 VGPRDFVSVRCAKRRGSTCVLAGMATDFGNMP[illegible]VIRAEHGPTCMVLHPLACSLSKIK 238
              | |:                        |:  ::  :::       ::  ||   :|    :    :|
Db      376 ARPADY---------------------DYL[illegible]KEFYSLRKYTPT [illegible]KAN 41[illegible]

Qy      239 LTWLLSIDLKGWLPKSIINQVLSQTQVDFANHL[illegible]KRLESH 278
                 | ::::    :  :|     :|     :  |||          :  |
Db      414 EPLLQAVELIPGMNESCKRKVPDDSPVDE---[illegible]SKWKRH 450

RESULT    8
US-09-328-869-2
; Sequence 2, Application US/09328869B
; Patent No  6168933
; GENERAL INFORMATION:
;  APPLICANT:        Kaser, Matthew, R.
;  APPLICANT:  Hillman, Jennifer, L.
;  APPLICANT:  Baughn, Mariah, R
;  TITLE OF INVENTION:  PHOSPHOLIPID TRANSFER PROTEIN
;  FILE REFERENCE:  PC-0003 US
;  CURRENT APPLICATION NUMBER: US/09/328,869
;  CURRENT FILING DATE:  1999-06-08
;  NUMBER OF SEQ ID NOS:        17
;  SOFTWARE:          PERL Program
; SEQ ID NO 2
;    LENGTH: 291
;    TYPE: PRT
;    ORGANISM: Homo sapiens
;    FEATURE: -
;    OTHER INFORMATION: 1772859CD1
US-09-328-869-2

  Query Match               6.7%;  Score 98.5  DB 4;  Length 291;
  Best Local Similarity   23.8%;  Pred. No. [illegible].012;
  Matches   34;  Conservative   26;  Mismatches   56;  Indels   27;  Gaps   4;

Qy      127 DQPMERLYEELVERMEAMCEWNPNVKEIKV[illegible]THELAAEAAGNLVG------ 180
             | | | ||: |  :  :|    :|: || |                   |  ::|          ||
Db       76 DVPAETLYDVLHD-IEYRKKWDSNVIE-----------TEDIARLTVNADVGYYSWRC 121

Qy      181 -----PRDFVSVRCAKRRGSTCVLAGMATDF[illegible]NMPEQKGVIRAEHGPTCMVLHPLAGSPS 2[illegible]
                  || :::|        |:   ::                 | :|  ::||       |   ::     :   |
Db      122 FKPLENR[illegible]VPAVS[illegible]GYL[illegible]--SIGPK 1[illegible]

Qy      2[illegible] K[illegible]TWL[illegible]SIDLKGWLPKSIINQ 258
                 :|:|    :|  ||  |||  ::|:
Db      180 SCVITYLAQVDPKGSLPKWVVNK 202

RESULT    9
```

```
US-09-198-956-2
; Sequence 2, Application US/09198956
; Patent No. 6165769
; GENERAL INFORMATION:
;  APPLICANT: Andersen, Lene N.
;  APPLICANT: Schulein, Martin
;  APPLICANT: Lange, Niels Erik K.
;  APPLICANT: Bjornvad, Mads E.
;  APPLICANT. Schnorr, Kirk
;  TITLE OF INVENTION: Pectin Degrading Enzymes From Bacillus
;  TITLE OF INVENTION: Licheniformis
;  FILE REFERENCE: 5377.200-US
;  CURRENT APPLICATION NUMBER: US/09/198,956
;  CURRENT FILING DATE: 1998-11-24
;  EARLIER APPLICATION NUMBER: [illegible]44/97
;  EARLIER FILING DATE: 1997-11-24
;  EARLIER APPLICATION NUMBER [illegible]
;  EARLIER FILING DATE: 1997-12-02
;  NUMBER OF SEQ ID NOS: 26
;  SOFTWARE: FastSEQ for Windows Version [illegible]
; SEQ ID NO 2
;    LENGTH: 494
;    TYPE: PRT
;    ORGANISM: Bacillus licheniformis
US 09 198-956-2

  Query Match               6.2%;  Score 9[illegible].5  DB 4;  Length 494;
  Best Local Similarity   22.3%;  Pred. No. [illegible];
  Matches   59;  Conservative   39;  Mismatches   1[illegible]9;  Indels   57;  Gaps   10;

Qy       52 VRRRSSL[illegible]SRLEETLYSDQELAYLQQGE[illegible]K 111
             :|   ||     |    |       :|:  |  |    :  :    |         :|  |   ::          |:  :
Db       40 IRLESS---SGLNITPAGNQDNAPLTAKQ[illegible]EK ------EERW[illegible]LDISDGKQFKIFNM 89

Qy      112 VVPDV[illegible]KLEVVVDQPMERL[illegible]NPNVK[illegible]KVLQKIGKDIEF[illegible]IH 168
                  |  ||:         :  : |      |:         |  :   ||        |      |  |    | |:
Db       90 ---DSGKIIIPAHYALSDNNPAVVYY[illegible]N--[illegible]KEEL  WN  [illegible]CADKDCNGDE[illegible]IY 137

Qy      159 ELAAEAAGNLVGPRDFVSVRCAK--------[illegible]--RRGS[illegible]CV 198
             :: :      :|      |     |: ||                                      : |:
Db      138 KIVSAQN[illegible]LA[illegible]LDCSCVKLAK[illegible]TGTIGG 197

Qy      199 LAGMATD[illegible]NMPEQKGVIRAEHGPTCMVL[illegible]NQ 2[illegible]
             | |          : | |   |   :   |   :|:      |:  :  |  ||           :  |     |    :|
Db      198 LLGKVVTVNN[illegible]ELKANIE[illegible]KLNN 2[illegible]

Qy      259 VLSQTQVDFANHLRKRLE SHPAS 281
             :    :|:     |  :  |  |    :|  ||
Db      257 IYFK[illegible]EA[illegible]NV[illegible]EKNLVIAHDAS 28[illegible]

RESULT  10
US-08-538-960-5
; Sequence 5, Application US/08538960
; Patent No. 5872230
;   GENERAL INFORMATION:
;      APPLICANT:  Stocco, Douglas M.
;      APPLICANT:  Clark, Barbara J.
;      TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR
;      TITLE OF INVENTION:  REGULATION OF STEROIDOGENESIS
;      NUMBER OF SEQUENCES:  14
;      CORRESPONDENCE ADDRESS:
;        ADDRESSEE:  Akin, Gump, Strauss, Hauer & Feld, L.L.P.
;        STREET:  1900 Frost Bank Plaza, 8[illegible] Congress Avenue
;        CITY:  Austin
;        STATE:  TX
;        COUNTRY:  U.S.A.
;        ZIP:  78701
;      COMPUTER READABLE FORM:
;        MEDIUM TYPE:  Floppy disk
;        COMPUTER:  IBM PC compatible
;        OPERATING SYSTEM:  PC DOS/MS-DOS
;
```

```
;        SOFTWARE: PatentIn Release #1.0, Version #[illegible]
;      CURRENT APPLICATION DATA:
;        APPLICATION NUMBER: US/08/[illegible]
;        FILING DATE: Concurrently Herewith
;        CLASSIFICATION: [illegible]
;      ATTORNEY/AGENT INFORMATION:
;        NAME: Mayfield, Denise L.
;        REGISTRATION NUMBER: [illegible]
;        REFERENCE/DOCKET NUMBER: [illegible]/EIM
;      TELECOMMUNICATION INFORMATION:
;        TELEPHONE: [illegible]/49[illegible]
;        TELEFAX: 512/499-[illegible]
;    INFORMATION FOR SEQ ID NO: 5:
;      SEQUENCE CHARACTERISTICS:
;        LENGTH: 19 amino acids
;        TYPE: amino acid
;        STRANDEDNESS: single
;        TOPOLOGY: linear
;      MOLECULE TYPE: peptide
US-08-538-960-5

  Query Match            [illegible]%;  Score [illegible];  DB [illegible];  Length [illegible];
  Best Local Similarity  [illegible]%;  Pred. No. [illegible];
  Matches   17;  Conservative  [illegible];  Mismatches  [illegible];  Indels  [illegible];  Gaps  [illegible];

Qy        194 GSICVLAGMA[illegible]GNM[illegible]EQ 212
              [illegible]
Db          1 GSICVLAGMA[illegible]GEM[illegible]EQ 1[illegible]

RESULT 11
US-08-659-254-5
; Sequence 5, Application US/08659254
; Patent No. 6194[illegible]
;  GENERAL INFORMATION:
;    APPLICANT: St[illegible], [illegible]ouglas M.
;    APPLICANT: Clark, [illegible] Barbara
;    TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR [illegible]
;    TITLE OF INVENTION: [illegible]STEROIDOGENESIS
;    NUMBER OF SEQUENCES: [illegible]
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE: Akin, Gump, Strauss, Hauer & Feld, [illegible]
;      STREET: 1900 Frost Bank Plaza, 816 Congress Ave[illegible]
;      CITY: Austin
;      STATE: TX
;      COUNTRY: U.S.A.
;      ZIP: 78701
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE: Floppy disk
;      COMPUTER: IBM PC compatible
;      OPERATING SYSTEM: PC-DOS/MS-DOS
;      SOFTWARE: PatentIn Release #1.0, Version #1.[illegible]
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER: US/08/659,254
;      FILING DATE: [illegible] JUN 199[illegible]
;      CLASSIFICATION: [illegible]
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER: US 08/538,9[illegible]
;      FILING DATE: 04 OCT 1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME: Mayfield, Denise L.
;      REGISTRATION NUMBER: [illegible]
;      REFERENCE/DOCKET NUMBER: [illegible]
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE: [illegible]
;      TELEFAX: 512/4[illegible]
;  [illegible]TION FOR SEQ ID NO: [illegible]
;    [illegible]CE CHARACTERISTICS:
;      [illegible]TH: 19 amino acids
;      [illegible]amino acid
;      [illegible]DEDNESS: single
;      [illegible]GY: linear
```

```
;    MOLECULE TYPE: peptide
US-[illegible]

  Query Match            [illegible]%;  Score [illegible];  DB [illegible];  Length [illegible];
  Best Local Similarity  [illegible]%;  Pred. No. [illegible];
  Matches  [illegible];  Conservative  [illegible];  Mismatches  [illegible];  Indels  [illegible];  Gaps  [illegible];

Qy        194 [illegible]VLAGMA[illegible]GNM[illegible]
              [illegible]
Db          1 [illegible]VLAGMA[illegible]GEM[illegible]

RESULT 12
US-09-464-486-4
; Sequence 4, Application US/09464486
; Patent No. 6[illegible]
;  GENERAL INFORMATION:
;    APPLICANT: Lawlor, Elizabeth [illegible]
;    TITLE OF INVENTION: Novel [illegible]
;    NUMBER OF SEQUENCES: [illegible]
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE: Dechert Price & Rhoads
;      STREET: 4000 Bell Atlantic Tower, 1717 Arch Street
;      CITY: Philadelphia
;      STATE: PA
;      COUNTRY: US
;      ZIP: [illegible]
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE: Diskette
;      COMPUTER: IBM Compatible
;      OPERATING SYSTEM: DOS
;      SOFTWARE: FastSEQ for Windows Version [illegible]
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER: US/09/464,486
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER: US/09/[illegible]
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME: Kirk, [illegible]
;      REGISTRATION NUMBER: [illegible]
;      REFERENCE/DOCKET NUMBER: [illegible]
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE: [illegible]
;      TELEFAX: [illegible]
;      TELEX:
;  INFORMATION FOR SEQ ID NO: 4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH: [illegible] amino acids
;      TYPE: amino acid
;      STRANDEDNESS: single
;      TOPOLOGY: linear
US-09-464-486-4

  Query Match            [illegible]%;  Score [illegible];  DB 4;  Length 4[illegible];
  Best Local Similarity  [illegible]%;  Pred. No. [illegible];
  Matches  [illegible];  Conservative  [illegible];  Mismatches  [illegible];  Indels  [illegible];  Gaps  [illegible];

Qy        [illegible] [illegible]AMLKA[illegible]NGEGWKK[illegible]NG[illegible]KVMSKVVE[illegible]V[illegible]KVF [illegible]
              [illegible]
Db        [illegible] [illegible]AMV[illegible]VKR[illegible]VVEN[illegible]VN[illegible]SVD[illegible]

Qy        [illegible] [illegible]VVV[illegible]MF[illegible]YEELVERMFAMGEWNEN [illegible]
              [illegible]
Db        [illegible] [illegible]GMKVHERK[illegible]VVV[illegible]YBAE[illegible]ACK[illegible]FA[illegible]

Qy        [illegible] [illegible]VKE[illegible]AAEAA[illegible]NEV[illegible]EVSVE[illegible]AKRE[illegible]VLAGMA[illegible]GNM [illegible]
              [illegible]
Db        [illegible] [illegible]VN[illegible]KEVE[illegible]ANE[illegible]AEA[illegible]YKEN[illegible]   A[illegible]AENV [illegible]
```

```
Qy     210 PEQKGVIRAEHGPTCMVLHPLACSPSKTKLTWLISI LKGWLPKSIINQVLSQTQVDFAN 269
                                         ||      :||     :|::::|  | :    :||
Db     260 EKEEAI--------------------TKAIDNIEII---IPEAMVNIELDRMVSEFAQ 294
            ::: :

Qy     270 HLRKR 274
            ::::
Db     295 RIQQQ 299


RESULT  13
US-09-414-664-4
; Sequence 4, Application US/09414664
; Patent No. 6242249
;  GENERAL INFORMATION:
;    APPLICANT:  Lawlor, Elizabeth J.
;    TITLE OF INVENTION:  No. 6242249el tic
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dechert Price & Rhoads
;      STREET:  4000 Bell Atlantic Tower, 1717 Arch Stre
;      CITY:  Philadelphia
;      STATE:  PA
;      COUNTRY:  US
;      ZIP:  19103
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/414,664
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/999,339
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Falk, Stephen T
;      REGISTRATION NUMBER:  36,795
;      REFERENCE/DOCKET NUMBER:  GM1005
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  215-994-2488
;      TELEFAX.  215-994-2222
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  402 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
  09-414-664-4

  Query Match             5.6%;  Score 8 ;  DB 4;  Length 402;
  Best Local Similarity   23.3%;  Pred. No. ;
  Matches   57;  Conservative   37;  Mismatches   71;  Indels   80;  Gaps   11;

Qy      64 EETLYSDQELAYLQQGEFAMQKALGILMNQTWKKPSQQDNGDKVMSKVVPDV-GKVF-- 120
            :||  || ||        :||:  :|| |.     .::    :|||  |           | |:
Db     101 QETELSDDEL  ---QEAIDHSLGHLAEMV VKEDCVVENGDIVNIIESGSVDCEEEEG 153

Qy     121 ------------------------- --R LEVVVDQPMEELYEEIVPEMEANCEWNPN 150
                                         :|||   | |    |||   : |       :
Db     154 CQAECYDLFICSGSFIPGFEEQLEGMKVDEEKLEVVVTEPEEYHAEELACK-EA--IFKIK 210

Qy     151 VKETKVLQKIGKDTEITHELAAEAAGNLVG  EPEVSVRCAKRRGSTCVLAGMAIDEGNM 209
            | |||   :          | :|| |||    | |      : :   | |:::          :|  | .
Db     211 VNEIKFKEVPELTDETANELDAEA  NEV  PVENIEPYELAPQK  -- - -  AIDAPNV 259

Qy     210 PEQKGVIRAEHGPTCMVLHPLACSPSKIKL WLISIDLKGWLPKSIINQVLSQTQVDFAN 269
                                         ||        :||      :|.... |  |  .    :||
Db     260 EKEEAI--------------------TKA DNIID----IPEAMVNIELDRMVSEFAQ 294
            ::: :
```

```
Qy     270 HLRKR 274
            ::::
Db     295 RIQQQ 299


RESULT  14
PCT-US95-10166-2
; Sequence 2, Application PC/TUS9510166
;  GENERAL INFORMATION:
;    APPLICANT:  Duke University
;    TITLE OF INVENTION:  USE OF ZRK PEPTIDE SEQUENCES IN
;    TITLE OF INVENTION:  CONTRACEPTION
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Banner & Allegretti, Ltd
;      STREET:  1001 G Street, N.W.
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20001
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0  Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/10166
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/289,1
;      FILING DATE:  11-AUG-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kagan, Sarah A.
;      REGISTRATION NUMBER:  32,141
;      REFERENCE/DOCKET NUMBER:  00240.16 80
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-508-9100
;      TELEFAX:  202-508-9299
;      TELEX:  97430 BBMB UT
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  600 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  YES
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Homo sapiens
PCT-US95-10166-2

  Query Match             5.6%;  Score 8 ;  DB 5;  Length 600;
  Best Local Similarity   21.1%;  Pred. No. 2. ;
  Matches   57;  Conservative   47;  Mismatches   80;  Indels   86;  Gaps   14;

Qy      31 ISQELNRRALGGPTIPSTWINQV-----   RKRSS   ------  - LLG- ---  ---- 60
            :| |:   .    |  ||  |  :  ::                :|||              |:|
Db     123 VSDEVQRGMPGNVIPCIRLGRLCPLENL AWQRRSEALEGWLLSWTTLLGLVLYISLAIR 182

Qy      61 SRLEETLY------SDQELAYLQQGEEAMQKALGILSNQEGWKK ISQQDNGDKVMSKVVP 114
            |::||  :         |  ||        ::      |::   |   :    |         |:  :  :
Db     183 KRVQETKEGNAFTEEDSELVVNYIAKKSFCRRAIELTESLGVSEEELQNKLEEVVIDENLL 242

Qy     115 DVGKVFRLEVVVDQPMEELYEELVTKMEAWTWN  NVKEEIKV---LQKIGKDTEITHEL 170
            ||         . . .       :||           |            ||:       : :   ||     |:
Db     243 ILGK -  -- LIAGEGKGIVTE                         TWN  EGKEVK LVAIKTLKEL THANKEL 288

Qy     171 AAEAA-    GN-----LVGPRDPVSVRCAKRRGSTCVLA    EPEL INPEEQWTVLYA 218
            ::       ||       |:|                       :  :     :  |::     :    ::     :
Db     289 LDEASVMASFGNPHVVRLLG---  -                -   --- ISMIST  VTIEKEGAVYKELKDKA 333
```

```
Qy      219 E HGP[illegible]MV[illegible]PLA[illegible]SP[illegible]KIKLIWL[illegible] 2[illegible]
                | |     |      |        |  | |
Db      334 [illegible]QHRSNCAF[illegible]NP     [illegible]QILLKEMV[illegible]IAL [illegible]

RESULT 19
US-09-464-483-2
; Sequence 2, Application US/09464483
; Patent No. 62280[illegible]
;  GENERAL INFORMATION:
;    APPLICANT: [illegible]
;    TITLE OF INVENTION: No. 62280[illegible]
;    NUMBER OF SEQUENCES: [illegible]
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE: Dechert Price & Rhoads
;      STREET: 4000 Bell Atlantic Tower, 17[illegible] Ar[illegible]
;      CITY: Philadelphia
;      STATE: PA
;      COUNTRY: US
;      ZIP: 1910[illegible]
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE: Diskette
;      COMPUTER: IBM Compatible
;      OPERATING SYSTEM: DOS
;      SOFTWARE: FastSEQ for Windows Version [illegible]
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER: US/09/464,483
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER: US/08/[illegible]
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME: Falk, Stephen [illegible]
;      REGISTRATION NUMBER: [illegible]
;      REFERENCE/DOCKET NUMBER: GMI[illegible]
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE: 21[illegible] 994 2188
;      TELEFAX: 215 994 [illegible]
;      TELEX:
;  INFORMATION FOR SEQ ID NO: 2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH: 52[illegible] amino acids
;      TYPE: amino acid
;      STRANDEDNESS: single
;      TOPOLOGY: linear
[illegible]9-464-483-2

  Query Match            [illegible]%;  Score 81;  DB 1;  Length [illegible];
  Best Local Similarity  [illegible]%;  Pred. No. [illegible];
  Matches  5[illegible];  Conservative  35;  Mismatches  [illegible];  Indels  [illegible];  Gaps  [illegible];

Qy       64 EEILYSDQELAY[illegible]QQGI[illegible]AMQKAL[illegible]II[illegible]SNQE[illegible]WKKE[illegible]Q[illegible]EN[illegible]KVM[illegible]KVVE[illegible] [illegible]KV[illegible] [illegible]
              :||   || ||                ||  :|: |                 ::     :|
Db      129 QEIEESDEEL         [illegible]ATD[illegible]SL[illegible]HE[illegible]AEMV VKE[illegible]VVENE[illegible] [illegible]

Qy      121                                              EIEVVV[illegible]PME[illegible] [illegible]
                                                          :||
Db      182 GQAE[illegible]YDLEE[illegible]SE[illegible]EIE[illegible]EEQLEEMKV[illegible]EEKIVVV[illegible]EEE[illegible]GAE[illegible]IA[illegible] [illegible]

Qy      151 VK[illegible]IKVLQKI[illegible]K[illegible]EI[illegible]EIAAEAAGN[illegible]V[illegible] [illegible]EV[illegible]V[illegible]AK[illegible] [illegible]NM[illegible]
             |  :||        :         |     |  |||    |  |
Db      239 VNEIKPKEVPEE[illegible]EIAN[illegible]DAEA  NIV[illegible]VKENE[illegible]KE[illegible]AL[illegible]E              [illegible]AENV [illegible]

Qy      210 PE[illegible]K[illegible]VIRAPH[illegible] [illegible]MV [illegible]PLA[illegible]SP[illegible]KIKLIWL[illegible]K[illegible]WL[illegible]K[illegible]IN[illegible]V[illegible] [illegible]AN [illegible]
              :    :                        |:                 :  :
Db          [illegible]KEEAI-                   [illegible]EAIDN[illegible]     [illegible]AMVN[illegible] [illegible]NV[illegible]A[illegible]

          [illegible] 272
          :
          Q [illegible]25
```

[illegible]